REHABILITATION OF SPORTS INJURIES: SCIENTIFIC BASIS

VOLUME X OF THE ENCYLOPAEDIA OF SPORTS MEDICINE
AN IOC MEDICAL COMMITTEE PUBLICATION

IN COLLABORATION WITH THE
INTERNATIONAL FEDERATION OF SPORTS MEDICINE

EDITED BY
WALTER R. FRONTERA
REHABILITATION OF SPORTS INJURIES:
SCIENTIFIC BASIS
IOC MEDICAL COMMISSION

SUB-COMMISSION ON PUBLICATIONS IN THE SPORT SCIENCES

Howard G. Knuttgen PhD (Co-ordinator)
Boston, Massachusetts, USA

Harm Kuipers MD, PhD
Maastricht, The Netherlands

Per A.F.H. Renström MD, PhD
Stockholm, Sweden
Contents

List of Contributors, vi
Forewords, viii
Preface, ix
Part 1: Epidemiology and Pathology
1 Epidemiology of Sports Injuries: Implications for Rehabilitation, 3
   W.R. Frontera
2 Pathophysiology of Injury, 10
   M.L. Schamblin and M.R. Safran
Part 2: Basic Science of Tissue Healing and Repair
3 Skeletal Muscle Regeneration After Injury: Cellular and Molecular Events, 35
   A.-X. Bigard and E. Fink
4 Tissue Healing and Repair: Tendons and Ligaments, 56
   B.W. Oakes
5 Tissue Healing and Repair: Bone and Cartilage, 99
   K.-M. Chan, H.C.L. Ho and C.W.C. Tong
Part 3: Practical Issues
6 Physiological and Performance Consequences of Training Cessation in Athletes: Detraining, 117
   I. Mujika and S. Padilla
7 Physiological and Functional Implications of Injury, 144
   C.J. Standaert and S.A. Herring
8 Psychological Factors in Sports Injury Rehabilitation, 160
   B.W. Brewer and A.E. Cornelius
Part 4: Clinical Rehabilitation Interventions
9 Pharmacological Agents and Acupuncture in Rehabilitation, 187
   J.K. Silver and J. Audette
10 Physical Modalities and Pain Management, 204
    J.M. Press, C.T. Plastaras and S.L. Wiesner
11 Flexibility and Joint Range of Motion, 232
    M. Schwellnus
12 Strength and Endurance, 258
    G. Grimby and R. Thomeé
13 Proprioception and Coordination, 274
    J.J. González Iturri
14 Functional Rehabilitation and Return to Training and Competition, 288
    W.B. Kibler and T.J. Chandler
15 Orthoses in the Prevention and Rehabilitation of Injuries, 301
    W. Micheo and A. Esquenazi
Index, 317
List of Contributors

J. AUDETTE MD, Instructor, Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston MA, USA

A.-X. BIGARD MD, Department of Human Factors, Centre de Recherches du Service de Santé des Armées, BP 87 La Tronche, 38702, France

B.W. BREWER PhD, Associate Professor, Department of Psychology, Springfield College, Springfield, Massachusetts 01109, USA

K.-M. CHAN MD, Chair Professor and Chief of Service, Hong Kong Centre of Sports Medicine & Sports Science, Department of Orthopaedics & Traumatology, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong

T.J. CHANDLER EdD, Associate Professor, Exercise Science, Sport, and Recreation, Marshall University, Huntington WV, USA

A.E. CORNELIUS PhD, Center for Performance Enhancement and Applied Research, Department of Psychology, Springfield College, 263 Alden Street, Springfield, MA 01109, USA

A. ESQUENAZI MD, Director, Gait and Motion Analysis Laboratory Moss Rehabilitation Hospital, Department of Physical Medicine and Rehabilitation, Jefferson College School of Medicine and Department of Bioengineering, Drexel University, Philadelphia PA, USA

E. FINK PhD, Department of Human Factors, Centre de Recherches du Service de Santé des Armées, BP87, La Tronche, 38702, France

W.R. FRONTERA MD, PhD, Chairman, Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston MA, USA

J.J. GONZÁLEZ ITURRI MD, Department of Physical Medicine and Rehabilitation, University of Navarra, 31007 Pamplona, Spain

G. GRIMBY MD, PhD, Professor Emeritus, Department of Rehabilitation Medicine, Göteborg University, Göteborg, Sweden

S.A. HERRING MD, Clinical Professor, Departments of Orthopedics and Rehabilitation Medicine, University of Washington, Seattle, Washington, USA

H.C.L. HO MBChB, Department of Orthopaedics and Traumatology, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong

W.B. KIBLER MD, Medical Director, Lexington Sports Medicine Center, Lexington, KY 40504, USA

W. MICHEO MD, University of Puerto Rico, Medical Sciences Campus, School of Medicine, Department of Physical Medicine, Rehabilitation and Sports Medicine, San Juan PR 00936-5067

I. MUJIKA PhD, Department of Research and Development, Medical Services, Athletic Club of Bilbao, Basque Country, Spain

B.W. OAKES MD, Associate Professor, Department of Anatomy and Cell Biology, Faculty of Medicine, Monash University, Clayton 3168, Melbourne, Australia
S. PADILLA MD, PhD, Department of Research and Development, Medical Services, Athletic Club of Bilbao, Basque Country, Spain

C.T. PLASTARAS MD, Rehabilitation Institute of Chicago, Center for Spine, Sports and Occupational Rehabilitation, 1030 N. Clark Street, Chicago, IL 60610, USA

J.M. PRESS MD, Rehabilitation Institute of Chicago, Center for Spine, Sports and Occupational Rehabilitation, 1030 N. Clark Street, Chicago, IL 60610, USA

M.R. SAFRAN MD, Co-Director Sports Medicine, Department of Orthopaedic Surgery, University of California San Francisco, San Francisco, California 94143, USA

M.L. SCHAUMBLEDIN MD, Department of Orthopaedic Surgery, University of California, Irvine, California, USA

M. SCHWELLNUS MD, PhD, Associate Professor, Sports Science Institute of South Africa, The University of Cape Town, Cape Town, South Africa

J.K. SILVER MD, Assistant Professor, Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston, MA, USA

C.J. STANDAERT MD, Clinical Assistant Professor, Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, USA

R. THOMEÉ PhD, Department of Rehabilitation Medicine, Göteborg University, Göteborg, Sweden

C.W.C. TONG MBChB (Hons), Department of Orthopaedics and Traumatology, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong

S.L. WIESNER MD, Chief, Occupational Health Department, The Permanente Medical Group, 280 West MacArthur Boulevard, Oakland, California 94611-5693, USA
The sports medical care of athletes is often wrongly assumed to comprise simply the immediate treatment of injuries and of systemic medical problems. The considerable time and the extensive effort devoted by medical and allied health personnel, both to the prevention of injuries and to the rehabilitation of athletes from injuries which curtail or prevent training and competition, are frequently overlooked.

This new volume in the Encyclopaedia of Sports Medicine series addresses all of the important issues related to the rehabilitation of the injured athlete. Dr Frontera and his team of expert contributing authors present the cutting edge of knowledge relative to the basic science and accompanying practical considerations regarding tissue injury and repair.

This volume provides an excellent complement to the volumes already published. The series now includes the preparation of an athlete for competition, the prevention of sports injuries, the immediate treatment of injuries, and the rehabilitation that must occur to bring an athlete back to training and competition.

My congratulations go to the editor and authors for their excellent work and to the IOC Medical Commission for providing this admirable contribution to sports medicine literature.

Dr Jacques ROGGE  
IOC President

The general aim of the Encyclopaedia of Sports Medicine series is to present the latest and most authoritative information available relative to a broad range of topic areas included under the rubric of Sports Medicine. The earlier volumes of the Encyclopaedia series have addressed a wide variety of areas of interest relative to both sports medicine and the sport sciences. Following the general interest publication of Vol. I, The Olympic Book of Sports Medicine, succeeding volumes were devoted to the more definitive topics of endurance, strength and power, prevention and treatment of injuries, the child and adolescent athlete, sports nutrition, women in sport, and biomechanics.

The publication of this volume further reinforces the intense interest that the IOC Medical Commission has in the health and welfare of the athletes of the world. Not only does optimal rehabilitation assist in returning an injured athlete to training and competition, but a carefully administered programme of rehabilitation serves to prevent the recurrence of the same injury or the occurrence of additional injuries. A high-quality programme of rehabilitation is of importance to all athletes, their coaches, and the teams and nations that they represent.

This volume will stand for many years as the most comprehensive and authoritative reference on sports injury rehabilitation available both for clinicians and sports scientists. I extend both my appreciation and my congratulations to Dr Frontera and each of the contributing authors.

Princes Alexandre de MERODE  
Chairman, IOC Medical Commission
Conceptual framework

Rehabilitation is, by definition, the restoration of optimal form (anatomy) and function (physiology). It is a process designed to minimize the loss associated with acute injury or chronic disease, to promote recovery, and to maximize functional capacity, fitness and performance. The process of rehabilitation should start as early as possible after an injury and form a continuum with other therapeutic interventions such as the use of pharmacological agents. It can also start before or immediately after surgery when an injury requires a surgical intervention. The rehabilitation of the injured athlete is managed by a multidisciplinary team with a physician functioning as the leader and coordinator of care. The team includes, but is not limited to, athletic trainers, physiotherapists, psychologists, and nutritionists. The rehabilitation team works closely with the athlete and the coach to establish the rehabilitation goals, to discuss the progress resulting from the various interventions, and to establish the time frame for the return of the athletes to training and competition.

Injuries during sports competitions may result from high forces during actions or movements inherent to the sport. The rehabilitation plan must take into account the fact that the objective of the patient (the athlete) is to return to the same activity and environment in which the injury occurred. Functional capacity after rehabilitation should be the same, if not better, than before injury since avoiding the conditions associated with the injury is not, in many cases, an alternative.

The sequence of events resulting from a sports-related injury that may lead to a reduction or inability to perform in sports can be framed using a disability model widely used in the field of rehabilitation medicine (Fig. 1). In this context, the ultimate goal of the rehabilitation process is to limit the extent of the injury, reduce or reverse the impairment and functional loss, and prevent, correct or eliminate altogether the disability.

From a clinical perspective it is possible to divide the rehabilitation process into three phases. The goals during the initial phase of the rehabilitation process include limitation of tissue damage, pain relief, control of the inflammatory response to injury, and protection of the affected anatomical area. The pathological events that take place immediately after the injury could lead to impairments such as muscle atrophy and weakness and limitation in the joint range of motion. These impairments result in functional

<table>
<thead>
<tr>
<th>Pathology (injury)</th>
<th>Impairment</th>
<th>Functional loss</th>
<th>Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td>Contracture</td>
<td>Inability to run, jump</td>
<td>Inability to compete in sports</td>
</tr>
<tr>
<td>Sprain</td>
<td>Muscle atrophy and weakness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fracture</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
losses, for example, inability to jump or lift an
object. The extent of the functional loss may be
influenced by the nature and timing of the thera-
peutic and rehabilitative intervention during the
initial phase of the injury. If functional losses are
severe or become permanent, the athlete now
with a disability may be unable to participate in
his/her sport.

The goals during the second phase of rehabili-
tation include the limitation of the impairment
and the recovery from the functional losses. A
number of physical modalities are used to en-
hance tissue healing. Exercise to regain flexibility,
strength, endurance, balance, and coordination
become the central component of the interven-
tion. To the extent that these impairments and
functional losses were minimized by early inter-
vention, progress in this phase can be accelerated.

The final phase of rehabilitation represents the
start of the conditioning process needed to return
to sports training and competition. Understand-
ing the demands of the particular sport becomes
essential as well as communication with the
couch. This phase also represents an opportunity
to identify and correct risk factors, thus reducing
the possibility of re-injury. The use of orthotic
devices to support musculoskeletal function and
the correction of muscle imbalances and inflexi-
bility in uninjured areas should receive the atten-
tion of the rehabilitation team.

Structure of this volume
This volume contains a total of 15 chapters
divided into four sections. The first section cov-
vers relevant basic concepts of the epidemiology
and pathology of sports injuries. The implica-
tions of the patterns of sports injury for rehabili-
tation are discussed and the physiological and
acellular response to tissue injury reviewed in
detail. The second section contains three chap-
ters on the basic science of tissue healing and
repair of the five most frequently injured tissues
in sports: muscles, tendons, ligaments, bones,
and/or cartilage. Section three includes three
chapters on practical issues of great significance
to the outcome of the rehabilitation process.
The treatment and rehabilitation of injuries in
athletes requires, in many cases, the reduction
or complete cessation of training. Injuries and
detraining alter basic physiological mechanisms
and the functional capacity of the athlete. These
effects must be taken into consideration espe-
cially at the beginning of the second phase of the
rehabilitation process. It is a common mistake to
consider the physical rehabilitation of the athlete
disconnected from the psychological recovery.
The last chapter of this section addresses relevant
emotional and psychological aspects of sports
rehabilitation.

The last section of the book includes seven
chapters that discuss the most commonly used
interventions in clinical rehabilitation. The use
of pharmacological agents, physical modalities,
and the various types of therapeutic exercise are
all discussed in detail. Particular attention is
given to the use of orthotic devices and to func-
tional rehabilitation and issues related to the
return to training and competition.

All authors have made a serious attempt to
summarize the relevant scientific literature. It is
our interest to discuss the evidence, if any, that
supports current rehabilitative strategies.

Acknowledgements
I would like to thank all authors for their time
and excellent contributions to this volume. I
also wish to express special thanks to Professor
Howard Knutten, Chair of the Sub-commission
on Publications of the IOC Medical Commission
for his guidance. Final thanks to the IOC Medical
Commission and the International Federation
of Sports Medicine for having established this
important Encyclopedia of Sports Medicine.

Walter R. Frontera
Boston, Massachusetts
Chapter 1

Epidemiology of Sports Injuries: Implications for Rehabilitation

WALTER R. FRONTERA

Introduction

The study of the relationships among the various factors determining the frequency and distribution of diseases and/or injuries in a human community is known as epidemiology. The basic elements of epidemiology have been applied to the study of a frequent, albeit unintended, consequence of the practice of sport; i.e. injuries to the musculoskeletal system.

Understanding the incidence and prevalence of injuries based on variables such as type and nature of the injury, age group, nature of the sport, gender and time since the onset of symptoms, among others, has contributed to the development of programmes aimed at the prevention and treatment of injured athletes (Walter & Hart 1990). Most importantly, these studies have resulted in the identification of risk factors for sports injuries (Macera et al. 1989) and modifications in the competitive rules in various sports. Although research in epidemiology has proved essential for the development of preventive and therapeutic interventions, the relationship between the epidemiology of sports injuries and the process of rehabilitation of the injured athlete has not received similar attention.

Some authors have proposed a model that uses epidemiological information to generate preventive strategies (Van Mechelen 1993). The sequence of prevention is illustrated in Fig. 1.1. In this model, the identification of the problem and description using epidemiological outcomes leads to the study of the mechanisms of injury and the naming and grouping of risk factors. Based on that information, preventive measures...
designed to reduce the risk and/or severity of the injuries are designed and implemented. Finally, the measures are evaluated by repeating the description of the problem (step 1) after the intervention. We suggest that this useful model could be expanded by including, as part of the intervention strategies, effective rehabilitation programmes that contribute to symptom resolution, limit functional losses, and restore physiological function and performance.

From a clinical point of view, an analysis of sports injuries by pattern, type, incidence and severity, together with an improved understanding of the physiological losses associated with these injuries, could help us design better rehabilitative interventions. Further, this knowledge could help us explain the extent to which the lack of effective rehabilitation itself becomes a risk factor predisposing injured athletes to the recurrence of an existing injury or to new injuries in a different, but related, anatomical area. For example, a high incidence of chronic injuries could indicate that proper rehabilitation did not follow the treatment of the symptoms in the acute inflammatory phase. It should be understood that, for the competitive athlete, resolution of the acute symptoms, such as pain, and clinical signs, such as swelling, is not the goal of the sports medicine practitioner. Restoration of form, and more importantly, function after resolution of the symptoms is necessary for optimal sports performance. Further, as our understanding of the physiological losses associated with the most common sports injuries improves, it will be possible to anticipate the functional deficits resulting from those injuries. Thus, the implementation of appropriate rehabilitation programmes will be feasible.

The purpose of this chapter is to illustrate how data on the pattern of injuries in various sports populations can help us restore form and function after injury. It is not the author’s intention to present an exhaustive and critical review of the literature on the epidemiology of sports injuries but to interpret some existing data in the context of the goals of a standard rehabilitation programme. It will suffice to demonstrate a reasonable association between the two. Further, most of the observations included in this chapter are descriptive of a sports injury clinic located in a sports training centre and not in a hospital or medical centre. The descriptive nature of these observations limits the extent to which we can draw conclusions and only allows us to make preliminary observations and speculations. Finally, it is not intended to present an analysis of the incidence of sports injuries with considerations of the population at risk or the difference in exposure (hours during which an athlete risks injury) (Wallace 1988). These factors appear to be less relevant in clinical rehabilitation.

Patterns of sports injuries

It is common to examine the distribution of injuries in relation to other variables of interest like age group, type of injury (traumatic vs. overuse), time since onset of symptoms, whether the injury occurred during training or competition, anatomical area, specific diagnosis and severity of injury. These variables are of significant interest when rehabilitation is our main focus. Let us examine briefly the influence of these factors on our rehabilitation strategies.

Type of injury (traumatic vs. overuse)

Roughly 45–60% of all injuries treated in a sports medicine clinic can be classified as overuse injuries. This is particularly true in sports like gymnastics (Fig. 1.2) where soft tissues and joints are subject to unusual positions and stresses. Risk factors for overuse injuries include muscle weakness, muscle strength imbalance and anatomical misalignment (Knapik et al. 1991). An examination of these risk factors suggests that properly designed rehabilitation programmes and the use of rehabilitation devices such as foot orthotics could contribute to a reduction in the incidence and prevalence of overuse sports injuries.

Clearly, the situation may be different when the analysis is restricted to clinical encounters.
in an emergency department. Many traumatic injuries (Fig. 1.3) are more acute and severe, resulting frequently in immobilization and prolonged rehabilitation.

Fig. 1.2 Repetitive stress associated with high-volume training results in overuse injuries in many athletes in various sports including gymnastics. M.O. Huilan at Atlanta, 1996. (© Allsport, Doug Pensinger, 1996.)

Fig. 1.3 Falls in cycling can result in significant traumatic injuries.

Time from onset

Epidemiological studies show that, when the time from onset of symptoms is considered, most (70.2% of a total of 1650) injuries treated in an ambulatory sports medicine clinic are chronic in nature (Frontera et al. 1994). In other words, the time between the onset of symptoms and the evaluation in the clinic is longer than 2 weeks. There are several potential reasons or scenarios that could explain this observation. As mentioned above, an insidious onset in the absence of obvious trauma with a slow progression of the injury could result in a delay in making a clinic appointment to get the signs and symptoms evaluated. This is typical of the overuse injuries that have become so prevalent in sports like running, swimming and baseball.

Another possibility could be that the symptoms are not well defined, at least initially, making the diagnosis by a physician difficult. It is also conceivable that the incorrect therapeutic modality or rehabilitation intervention was chosen to initiate treatment or that the patient did not complete the treatment as prescribed by the physician for other reasons. These two situations could prolong the acute stage resulting in the persistence of symptoms. Moreover, many injuries occur during training sessions when health care professionals may not be available to immediately treat...
the injured athlete. The athlete may delay seeking help or may initiate treatment him/herself if the symptoms are not severe or disabling, and tissue damage may increase in the absence of acute therapeutic intervention. Finally, it is also possible that the correct treatment was applied resulting in resolution of the symptoms but that proper rehabilitation of impairments and functional losses did not follow the therapeutic interventions. In other words, the athlete is allowed back into training and competition based on the absence of pain and inflammation but not on the recovery of strength, flexibility or endurance needed for successful performance in sports.

In the absence of appropriate rehabilitation, acute, subacute or chronic injuries frequently result in significant physiological and functional losses that affect the affected anatomical area (and adjacent tissues and joints) at risk for reinjury. Recovery from these losses becomes one of the most important goals of the rehabilitation programme. The extent of these losses is illustrated by a clinical epidemiology study published by Holder-Powell and Rutherford in 1999. These authors evaluated the strength of various muscle groups in asymptomatic subjects with history of a sports injury. The injuries occurred an average of 9.7 years before the evaluation (mean = 9.7 ± 11 years).

The most important observation in that study was a decrement of concentric, eccentric and static strength in the knee extensor muscles of the injured limb many years after the injury (Fig. 1.4). This was the case even when the muscle group was distant from the injured area. In other words, the authors of the study observed weakness of the knee extensors in patients with injuries such as fractures of the leg and sprains of the ankle ligaments. The degree of weakness present in the hamstrings, on the other hand, was minor in some cases and non-existent in others, suggesting that the rehabilitation approach must be muscle specific.

In another study, Croisier et al. (2002) demonstrated that athletes with strains of the hamstrings had significant strength deficits. More importantly, these investigators showed that if those that have significant deficits complete a rehabilitation programme, in this study consisting of isokinetic strengthening concentric and eccentric exercises, muscle weakness is reversed. Further, the incidence of postrehabilitation injuries in the 12 months following return to their sport was zero.

Anatomical distribution of injuries

When the incidence of sports injuries is analysed by anatomical region, the most frequently injured areas are the knee (Fig. 1.5), shoulder and ankle (Garrick 1985; DeHaven & Lintner 1986; Frontera et al. 1994). Of course the anatomical...
distribution of injuries in a particular sport can be very specific. In other words, basketball players may suffer more injuries to the knee than to the shoulder but the situation in swimmers is the reverse.

Knowledge of the anatomical distribution of injuries in a particular sport is essential to develop a training programme that maximizes sport-specific conditioning and minimizes the risk of injury. Further, because deconditioning associated with rest could potentially affect muscle groups proximal and distal to the injured area, knowledge of this anatomical distribution of injuries by sport could be vital for the rehabilitation of the injured athlete. A well-planned rehabilitation programme should include exercises for the injured area as well as for those areas at risk of injury in the specific sports activity.

**Most frequent diagnoses**

Most sports injuries are relatively mild and do not require surgical intervention. Independent of the level of competition, the most frequent diagnoses are in descending order: tendonitis (or tendinosis), first degree strains (muscle tendon unit), first degree sprains (ligament and capsular injuries), patellofemoral pain and second degree sprains. The best course of action in these cases is appropriate conservative intervention to control symptoms such as pain and swelling, followed by comprehensive rehabilitation. The indications for surgery in these cases are few and rehabilitation becomes the most effective intervention when fast return to practice or competition and prevention of future injuries are the most important goals (DeHaven & Lintner 1986; Matheson *et al.* 1989).

**Severity of injury**

The severity of the injury can be judged by the nature of the diagnosis, the duration and nature of the treatment, the time lost from sports training or competition, and/or the presence and degree of permanent damage (Van Mechelen 1993). There is usually a positive correlation between severity, functional loss and the need for extended rehabilitation. Clearly, when the severity is high, longer periods of immobilization or rest are needed for tissue healing. As a result, larger physiological losses are experienced by the athlete and deconditioning of uninjured areas is more extensive. Under these conditions, it should be anticipated that rehabilitation will last longer.

**Rehabilitation and the preparticipation exam**

Every competitive athlete must undergo a preparticipation medical examination on a regular basis. The preparticipation exam is an ideal situation in which to: (i) treat existing medical conditions early before the competition; (ii) anticipate the health care needs of the athlete; (iii) educate the athlete and his/her coach regarding health issues such as vaccinations and prevention of disease and injury; and (iv) discuss topics such as doping in sports.

Another important element of the preparticipation exam is the identification of risk factors for medical conditions in general and for sports injuries in particular. The process of identifying risk factors can make use of the epidemiologic evidence published in the sports medicine literature. Findings such as joint contractures or reduced flexibility, muscle weakness and muscle strength asymmetry represent ideal opportunities to do ‘preventive rehabilitation’. The restoration of normal form and function in these cases does not necessarily follow a sports injury but may be important in the prevention of future injuries. In addition, rehabilitation may prove to be beneficial for sports performance because an enhanced level of flexibility, cardiovascular endurance and muscle strength and endurance, alone or in combination, are required in almost any sport.

**Health services in international competitions**

The study of the pattern of disease and injuries in international sports competitions can help the team physician make plans regarding, among
other things, the composition of the health care team travelling with sports delegations, the equipment and medical supplies necessary to deliver health care (including rehabilitation services) in an Olympic village, the most commonly used and permitted medications, and the therapy modalities needed for rehabilitation (Frontera et al. 1997).

As the number of athletes and competitions increase and new transportation methods facilitate travel, health professionals will be challenged to respond to the needs of the travelling athlete exposed to different environmental conditions, pathogens and demanding training regimes.

Although disorders of the respiratory and gastrointestinal tracts are very common among travelling athletes, the main cause of morbidity during an international competition is injuries to the musculoskeletal system (Fig. 1.6).

The five most common injuries or disorders affecting the musculoskeletal system (total number of diagnoses = 966) include: first degree strains (23.1%), tendinitis/tendinosis (18.7%), concussion (12.7%), myositis (10.9%) and first degree sprains (10%). All of these diagnoses could benefit from rehabilitative interventions to control the symptoms in the acute stage and to restore physiological and functional capacity in later stages. In fact, in the above study, typical rehabilitation interventions such as physical therapy modalities (cold packs, hot packs, ultrasound, transcutaneous electrical stimulation, massage, therapeutic exercises) were needed in 23.1–36.9% of the total number of cases.

It is important to note that, just like in the case of the sports injury clinic discussed above, the onset of symptoms preceded the competition in many cases. Thus many injuries could be classified as chronic in nature. High-performance athletes often continue to train and compete even in the presence of symptoms and signs of injury, and may delay proper treatment and rehabilitation of an injury to participate in important competitions. Thus, it is reasonable to speculate that some of the injured athletes did not receive appropriate rehabilitation after the initial insult. These two observations make the inclusion of rehabilitation services in precompetition assessment and plans for a sports delegation an absolute necessity.

**Conclusion**

The study of the epidemiology of sports injuries can be as valuable to rehabilitation as it is to prevention. Many injuries may occur because the rehabilitation of a previous injury was not complete. Understanding risk factors associated with sports injuries can help in the design of rehabilitation strategies resulting in a lower incidence and severity of injuries. Rehabilitation principles can be applied in a sports injury clinic but also as part of the health care services for a travelling team. The following chapters will discuss the scientific basis of current rehabilitation practices. Every sports medicine practitioner should be familiar with these principles and apply them in their work with athletes.

**References**


Chapter 2

Pathophysiology of Injury

MARK L. SCHAMBLIN AND MARC R. SAFRAN

Introduction

No matter what the age of an athlete, the level of competition or the sport, inflammation is likely to affect an individual at some point in their endeavours. Whether an injury is one of chronicity, related to repetitive movements, or one of acute onset, related to trauma, the detrimental effects on athletic performance are well documented. Too frequently the complexity of the inflammatory process is not fully understood and inflammation is treated as an unwanted hindrance to athletic performance, however, it is truly a complex network of vascular and cellular responses designed to facilitate the repair of traumatized tissue (Bryant 1977; Gamble 1988; Martinez-Hernandez 1988).

The development of an inflammatory reaction to an injury is complex, utilizing many of the body’s systems to mediate its purpose. The goal of any inflammatory reaction is to resolve the pathological insult and restore the anatomy to a level of physiological function identical or nearly identical to preinjury status. Ideally this can be accomplished by removing diseased or damaged tissue with the subsequent regeneration of normal anatomical tissue. However, this is often not the case. Too frequently the insult is far too great or perpetrated over too long a period, resulting in increased tissue destruction. This often leads to scar tissue formation that in turn may propagate a continued inflammatory reaction. A persistent inflammatory action, therefore, may be harmful to an individual’s athletic performance, as well as to the individual.

Although in the recent past, understanding of the mediators of inflammation has vastly improved, many factors responsible for induction, regulation and resolution remain indefinable. This lack of understanding remains an elusive cornerstone in treatment for both the physician and the athlete. The purpose of this chapter is to provide an understanding of our current knowledge of the complex nature of the pathophysiological mechanisms, which function to mediate a host’s response to tissue injury. This knowledge is then utilized in later chapters to understand specific tissue responses to injury as related to the athlete.

Cytokines

Inflammation may be seen in a variety of circumstances that affect the human body. It may occur as a defensive response to foreign material or as a response to mechanical trauma, toxins or in the face of abnormalities such as neoplasia. The accumulation and activation of leucocytes seems to play an essential role in nearly all forms of inflammation. Following the influx of leucocytes in the acute phase of inflammation, its propagation and amplification is mediated by both humoral and cellular components of the immune system.

Cytokines are cellular proteins that are the mediators of physiological activity, including the inflammatory process. There are proinflammatory and anti-inflammatory cytokines, which modulate their effect by binding to receptors
on target cells. Through this interaction the up-regulation or down-regulation of cellular activity may be propagated. In the acute phase response of inflammation, cytokines appear to be responsible for a myriad of physiological modifications occurring both locally at the site of the pathology and also at regions distant from the insult (Rosenberg & Gallin 1993).

The functions of cytokines are varied. A given cytokine may initiate and regulate cellular activities in numerous cells simultaneously. At the same time, more than one cytokine may induce a particular biological activity (Table 2.1). Interleukin 1 (IL-1), a common cytokine seen in inflammation, acts on virtually all leucocytes, endothelial cells, monocytes and hepatocytes to up-regulate the expression of adhesion molecules, cytokines and arachidonic acid metabolites (Rosenberg & Gallin 1993). It results in neutrophil accumulation, fibroblast proliferation, angiogenesis, acute phase protein synthesis and metabolic alterations such as fever. Similar activity is seen in other proinflammatory cytokines, such as tumor necrosis factor alpha (TNF-α), IL-6, IL-4, IL-10 and transforming growth factor beta (TGF-β). An excellent example of this biological redundancy can be seen with IL-1 and TNF-α. Both of these cytokines result in the up-regulation of adhesion molecules, accumulation of leucocytes, protein synthesis and angiogenesis (Rosenberg & Gallin 1993; Bemelmans et al. 1996).

Local cytokines react with their receptors in an autocrine, paracrine and endocrine fashion (Fig. 2.1). The autocrine pathway allows for the amplification of the cytokine-induced inflammatory process. The paracrine pathway allows cytokines to influence cells in the local environment, leading to the accumulation of inflammatory cells. The induction of acute phase protein synthesis in the hepatocytes is an example of an endocrine mechanism of peripherally circulating cytokines (Akira et al. 1993; Dinarello 1996). By utilizing these three mechanisms of action, the cytokine is able to alter the local tissue as well as the acute phase response in the face of inflammation.

The regulation of cytokine function is of critical importance. Secondary to a cytokine’s powerful ability to modify biological behaviour, the body has evolved numerous cellular and molecular mechanisms to control their activity. The predominant mechanism of regulation occurs at the level of gene transcription. Cytokines are not stored, but rather created de novo following cellular activation. Antigen stimulation leads to increased transcription within 30 min. A steady-state level is seen in anywhere from 4 to 8 h (McDonald et al. 1993; Jain et al. 1995; Serhan 1997). With the cessation of the stimulus, return to the baseline will usually occur in 24 h. A second form of regulation is the conversion of an inactive form of the cytokine to the active form. This mechanism is seen in the conversion of a procytokine within the cytosol to the cytokine IL-1β, as well as the formation of TNF-α, which is formed in its active state but limited to the cell membrane. Cleavage of the membrane-bound TNF-α by converting enzymes facilitates secretion.

The ability to down-regulate the inflammatory process is as important as the ability to initiate it. Chronic inflammation or failure to control an
<table>
<thead>
<tr>
<th>Cytokine</th>
<th>Cell source</th>
<th>Cell target</th>
<th>Biological activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1α</td>
<td>Monocytes</td>
<td>All cells</td>
<td>Up-regulation of adhesion molecule expression</td>
</tr>
<tr>
<td></td>
<td>Macrophages</td>
<td></td>
<td>Macroage emigrations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Acute phase protein synthesis</td>
</tr>
<tr>
<td>IL-1β</td>
<td>T-cells</td>
<td>T-cells</td>
<td>T-cell activation and proliferation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B-cells</td>
<td>Enhanced monocyte and macroage cytolytic activity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td>IL-2</td>
<td>T-cells</td>
<td>T-cells</td>
<td>Stimulation of haematopoietic progenitors</td>
</tr>
<tr>
<td></td>
<td>Mast cells</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NK cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mast cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IL-3</td>
<td>T-cells</td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mast cells</td>
<td>Basophils</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NK cells</td>
<td>B-cells</td>
<td></td>
</tr>
<tr>
<td>IL-4</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mast cells</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Basophils</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mast cells</td>
<td></td>
</tr>
<tr>
<td>IL-5</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mast cells</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NK cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mast cells</td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IL-6</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Macrophages</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IL-7</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Macrophages</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IL-8</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Macrophages</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IL-9</td>
<td>T-cells</td>
<td>T-cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Macrophages</td>
<td>Monocytes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td>Macrophages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eosinophils</td>
<td></td>
</tr>
<tr>
<td>IFN-α</td>
<td>All cells</td>
<td>All cells</td>
<td>Fever, anorexia and proinflammatory cytokine production</td>
</tr>
<tr>
<td>IFN-β</td>
<td>T-cells</td>
<td>All cells</td>
<td>Cell cytotoxicity</td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGF-α</td>
<td>Most cell types</td>
<td>Most cell types</td>
<td>Down-regulates T-cell and macroage responses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stimulates angiogenesis</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>T-cells</td>
<td>All cells</td>
<td>Regulates macroage and NK cell activation</td>
</tr>
<tr>
<td></td>
<td>NK cells</td>
<td></td>
<td>T-cell differentiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Immunoglobulin production by B-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IFN, interferon; IL, interleukin; NK, natural killer; TGF, transforming growth factor; TNF, tumour necrosis factor.
inflammatory process may lead to host tissue damage. There are several mechanisms utilized in the down-regulation of inflammation. These include production of activated complement inhibitors, apoptosis of inflammatory cells and production of anti-inflammatory cytokines such as IL-4, IL-10, IL-13 and TGF-β (Feng et al. 1996). IL-4 and IL-10, perhaps the best known of the anti-inflammatory cytokines, appear to mediate an anti-inflammatory effect on the T-cell predominantly, but also B-lymphocytes, mast cells, basophils and endothelial cells, as well as a variety of others (Feng et al. 1996).

Cytokines are potent proteins in the initiation, propagation and regulation of the inflammatory process. In this regard they are not alone, as the body synthesizes various types of proteins to mediate these same functions. Among these are prostaglandins and leukotrienes, which will be discussed in the subsequent sections.

**Prostaglandins**

Along with cytokines, prostaglandins are amongst the best-defined mediators of the inflammatory response. Since their discovery in 1931, advances in their structure, function and physiological mechanisms have afforded an increased understanding of these molecules (Kurzok & Lieb 1931). Independent work by two groups demonstrated arachidonic acid conversion to prostaglandin E2 via the enzyme cyclooxygenase (Goldblatt 1933; Von Euler 1935; van der Pouw et al. 1995). From this finding it was thought that essential fatty acids served merely as a precursor to prostaglandin synthesis. This has subsequently been shown to be only one of the many functions of fatty acids, albeit an important one.

There are many ways of inducing prostaglandin synthesis that appear to be cell specific. In macrophages, prostaglandin E₂ (PGE₂) and thromboxane A₂ (TxA₂) are stimulated by the presence of antigen–antibody complexes (Poranova et al. 1996). Cytokine receptors on mast cells stimulate the synthesis and secretion of PGD₂ (Murakami et al. 1994). IL-1 and TNF-α stimulation of endothelial cells and fibroblasts leads to PGE₂ as well as PGI₂ production. The production of prostaglandins is accomplished by the breakdown of membrane phospholipids by phospholipase A₂ with the subsequent formation of arachidonic acid. Arachidonic acid is then converted to PGG₂ via cyclooxygenase 1 and 2. PGG₂ then may be converted into various prostaglandins by prostaglandin synthase (Fig. 2.2).

Investigations of the role of prostaglandins within the inflammatory cascade are extensive. In general, their function has been delineated by their injection into both animal and human subjects with subsequent monitoring of their effects. Another area of focus is the role of non-steroidal, anti-inflammatories in the regulation

---

**Fig. 2.2** The synthesis of prostaglandins and thromboxane A₂ from the membrane phospholipids, utilizing phospholipase A₂, cyclooxygenase and prostaglandin synthase.
of cyclooxygenase function. Throughout these studies, one thing has remained clear, the biological responses instigated by the presence of prostaglandins are wide ranging and varied in nature. In an attempt to simplify their function, their actions may be broken down into four general areas of the inflammatory process: fever, pain, oedema and leucocyte regulation.

Fever appears to be a complex neuroendocrine response to both infection and inflammation. A variety of proinflammatory mediators appear to serve as a stimulant to fever production. Various cytokines, which will be discussed in later sections, assist in the stimulation of the thermoregulatory centre by stimulating both the synthesis of the PGE family as well as mediating a direct effect on the thermoregulatory centre itself. Increased levels of PGE₂ have been demonstrated in the cerebral spinal fluid of febrile animals. This PGE₂ is most probably synthesized within the central nervous system as a result of bacterial-induced cytokine release. PGE₂ appears to modulate its effect on thermoregulatory centres within the hypothalamus; however, the specific receptor of action remains elusive (Feldberg & Saxena 1971; Milton & Wendlandt 1971). Cyclo-oxygenase inhibitors, specifically cyclooxygenase 2 inhibitors, given to human subjects modulate an antipyretic effect, lending further support to the role of prostaglandins in the induction of a febrile state.

In experimental models, the injection of prostaglandins in itself does not induce a painful response. However, in the presence of prostaglandins, particularly PGE₂ and PGI₂, the response to painful stimuli is greatly increased (Ferreira et al. 1978). It is unclear which of these two prostaglandins serves to accentuate the painful stimuli to the greatest extent or if different individuals respond to different stimuli in altered manners. In some experimental studies, PGE₂ appears to predominate while in others PGI₂ appears to be the prime mediator of an increase in the nociceptive response (Mnich et al. 1995; Plemons et al. 1996).

Similar to the response seen in pain, the injection of prostaglandins alone does not stimulate oedema. In states of inflammation, however, an increase in oedema is seen by the ability of prostaglandins to dilate the vasculature leading to increased blood supply to traumatized areas (Moncada et al. 1973; Wheeldon & Vardey 1993). In the presence of inflammation, increased vascular permeability is present secondary to the action of many proinflammatory mediators; this increased permeability coupled with increased blood flow stimulates excessive oedema. Experimental models also support this finding, as the injection of prostaglandins, specifically PGE₂ and PGI₂, coupled with bradykinin or platelet-aggregating factor (PAF) (powerful stimulants to increased permeability) stimulate oedematous states (Von Euler 1934).

In an apparent contradiction, the systemic administration of prostaglandins appears to mediate an anti-inflammatory effect (Kunkel et al. 1979; Fantone et al. 1980). Circulating prostaglandins (PGE₂ and PGD₂) inhibit neutrophils, monocytes and circulating T-lymphocytes. The inhibition of neutrophils is accomplished by inhibition of activation as measured by chemotaxis and superoxide production (Wedmore & Williams 1981). The inhibition of monocytes is mediated via the decreased production of various proinflammatory mediators such as TNF-α (Kunkel et al. 1986). T-lymphocyte inhibition is seen as a decrease in T-cell proliferation, a decrease in released cytokines and a decrease in the number of migrating T-cells (Goodwin et al. 1977; Shaw et al. 1988; Betz & Fox 1991; Trinchieri 1995). Through these three mechanisms, systemic prostaglandins appear to inhibit the inflammatory reaction, whereas, as discussed previously, the local accumulation of prostaglandins accentuates the inflammatory response.

As one can see, in recent years, the role of prostaglandins has been greatly elucidated and their involvement in the inflammatory reaction has been well documented. Secondary to their complex nature, however, it is highly probable that their involvement is even more significant than is currently known. Recent advances in understanding the nature of these molecules have led to improved pharmacological agents,
such as specific cyclooxygenase inhibitors in anti-inflammatory medications. Despite the success of these agents, one thing is clear—prostaglandins are merely a fraction of the known types of pro-inflammatory mediators that affect the human body in the face of pathology.

In addition to mediators such as cytokines and prostaglandins, the body’s immune system is composed of a highly complex series of cellular and plasma-derived components. The cellular component consists of a variety of cell types, each with a specific function. These cells interact in a well-orchestrated manner to improve the efficiency of the immune response. The cell types are variable and functions range from stimulating the aggregation of leucocytes to presenting foreign material to the actual destruction and removal of the offending pathogen. Some of the more important cellular components of the inflammatory process will be outlined in the following pages.

**Mast cells and basophils**

Several components of the immune system play a critical role in inflammatory reactions. Tissue mast cells and circulating basophils are haematopoietically derived cells, which express a variety of surface receptors that allow them to migrate to specific tissue locations, interact with cells and tissues, and respond to activation molecules. Both types of cells contain granules, which serve as storage for histamine, serotonin, cytokines and proteases. When IgE-sensitized mast cells or basophils are stimulated by either antigens or C3a and C5a (complement anaphylatoxins), the granules are released resulting in the secretion of these mediators. These in turn induce a reversible cell contraction in the endothelium, leading to the formation of gap junctions (Black et al. 1972; Arrang et al. 1983). This increases the permeability of the vasculature leading to increased tissue oedema. Both mast cells as well as basophils may be induced to release histamine and serotonin by other means, including physical stimuli such as trauma or proteins secreted by activated platelets and neutrophils.

**Macrophages**

Macrophages are present throughout a variety of host tissues (Fig. 2.3). These cells are able to react to abnormalities such as ischaemia or metabolic disturbance by initiating an inflammatory reaction. When local processes are insufficient and unable to remove the host tissue of the initiating stimuli, macrophages along with other cells can mobilize other forms of leucocytes (polymorphonuclear neutrophils in particular) by the activation of local endothelium and by the production of a variety of chemokines, cytokines and other lipid mediators of the inflammatory reaction (Table 2.2) (Bacon & Schall 1996). If the pathology continues and becomes one of chronicity, macrophages are able to up-regulate their microbicidal activity (Bell et al. 1994). When considering the role of macrophages in the initiation, propagation and resolution of inflammation, it is important to remember that each macrophage contains variable receptors on its membrane, allowing it to regulate the biosynthesis and secretion of substances in response to stimuli from the host tissue (Peterson et al. 1987).

As noted previously, macrophages interact with a variety of cells within the human body. These interactions are complex and reciprocal in nature. Non-haemopoietic cells, such as endothelium, are dramatically affected by the secretory products of the macrophage, and in turn have a profound effect on the ever-adapting macrophage itself.
In this way, one can see how the macrophage is an important component in the regulation of the inflammatory process.

**Neutrophils**

These cells maintain the ability to mobilize from the blood to the tissue with subsequent degranulation in a matter of seconds. Their major function in the inflammatory cascade is one of endocytosis (eating) or exocytosis (secreting) (Bainton 1980). In the normal adult human, a polymorphonuclear neutrophil is found in one of three environments: bone marrow, blood or tissues. The bone marrow is the site where proliferation and maturation occurs. Following the phase of proliferation and maturation, the neutrophils are released into the blood where they circulate for approximately 10 days. They then migrate into the tissues where they survive for approximately another 1–2 days. Their ultimate fate after this is unknown (Bainton 1980).

Four distinct populations of granules have been identified within neutrophils by cytochemical and cell-fractionation procedures (Borregaard & Cowland 1997). The azurophil granules contain myeloperoxidase (an antibacterial enzyme), lysozyme and lysosomal enzyme, as well as a variety of other agents (Table 2.3) (Klebanoff & Clark 1978). Specific granules by definition do not contain peroxidase (Cramer & Breton-Gorius 1987; Livesey et al. 1989; Mutasa 1989; Path et al. 1996). These granules contain numerous agents including lysozyme and lactoferrin (Bretz & Baggioni 1974). Gelatinase granules are subsets of specific granules, and are therefore peroxidase negative (Borregaard & Cowland 1997). They are named for their high content of gelatinase found in their granules (Borregaard et al. 1993; Kjelsen et al. 1993; Borregaard & Cowland 1997). Secretory vesicles are a group of vesicles that are easily mobilized to the surface; they are remarkable for the presence of alkaline phosphatase within the membrane as well as the presence of albumin (Borregarrd et al. 1990; Borregaard 1996).

The degranulation of neutrophils is mediated by the presence of an injury. The azurophil and secretory granules can be released independently (Williams & Morley 1973; Wright et al. 1977; Presentey 1984). However, depending on

<table>
<thead>
<tr>
<th>Cytokines</th>
<th>Multiple local and systemic host defence functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1</td>
<td>Multiple local and systemic host defence functions</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Antiviral, immune modulation</td>
</tr>
<tr>
<td>IFN-α/β</td>
<td>Inhibits proinflammatory cytokines</td>
</tr>
<tr>
<td>IL-6</td>
<td>Acute phase response</td>
</tr>
<tr>
<td>IL-10</td>
<td>Inhibits activation of macrophages and other cells</td>
</tr>
<tr>
<td>TGF-β</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complement proteins</th>
<th>Local opsonization and complement activation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most components</td>
<td></td>
</tr>
<tr>
<td>Coagulation factors</td>
<td>Initiation and regulation of clot formation</td>
</tr>
<tr>
<td>Adhesion and matrix molecules</td>
<td>Localization and migration</td>
</tr>
<tr>
<td></td>
<td>Modulates cellular interactions and phagocytosis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bioactive lipids</th>
<th>Mediators of inflammation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclooxygenase, lipoxygenase</td>
<td></td>
</tr>
<tr>
<td>Platelet-activating factor</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antimicrobial activity</th>
<th>Killing and stasis of microbial targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superoxide anion</td>
<td></td>
</tr>
<tr>
<td>Hydrogen peroxide</td>
<td></td>
</tr>
<tr>
<td>Nitric oxide</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.2 Macrophage-derived secretory products.**
the stimuli, concomitant release is required to accentuate the bactericidal effects. There appears to be a hierarchy in ability to mobilize granules (Borregaard et al. 1993). The hierarchy for mobilization for excretory function appears to be secretory vesicles, gelatinase granules, specific granules and finally azurophilic granules being the least likely to be mobilized (Borregaard 1996). When activated, the specific granules, gelatinase granules and the secretory vesicles bind to the plasma membrane via cytochrome b558 subunits. Their contents are readily released; however they lack the ability to generate reactive oxygen molecules without the contents of the azurophilic granules, specifically myeloperoxidase (Klebanoff & Clark 1978; Pryzwansky et al. 1979).

The clinical importance of proper neutrophil function can be seen in a variety of hereditary disease states such as acute myelogenous leukaemia, congenital dysgranulopoietic neutropenia or Chediak–Higashi syndrome (Bainton 1975; Bainton et al. 1977). In congenital dysgranulopoietic neutropenia there is a defective synthesis and degradation of azurophilic granules, an absence or marked deficiency of specific granules and autophagia. Patients with this disease suffer severe life-threatening infections. In Chediak–Higashi syndrome, a rare autosomal recessive disease, there is a presence of abnormally large inclusions within the neutrophil, which appear to be abnormal azurophilic granules (Davis & Douglas 1971; Ohashi et al. 1992). These patients demonstrate an increased susceptibility to infection.

**Eosinophils**

Eosinophils are a type of leucocyte identifiable by its bilobed nuclei and large eosinophilic granules. The large granules in the eosinophil contain peroxidase; however, this peroxidase is chemically different than the peroxidase found in neutrophils (Bujak & Root 1974). Eosinophil peroxidase appears to have no role in the bactericidal activity of eosinophils. The granules also contain a variety of proteins including major basic protein (MBP), an eosinophil cationic protein, which does appear to be cytotoxic to either parasites or mammalian cells. MBP is also responsible for the induction of histamine release by basophils and

---

**Table 2.3** Contents of neutrophil granules and vesicles.

<table>
<thead>
<tr>
<th>Azurophil granule</th>
<th>Specific granule</th>
<th>Gelatinase granule</th>
<th>Secretory vesicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cd63</td>
<td>Cd66</td>
<td>Cd11b</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>Cd68</td>
<td>Cd67</td>
<td>Cytochrome b558</td>
<td>Cytochrome b558</td>
</tr>
<tr>
<td>β-glycerophosphatase</td>
<td>Cytochrome b558</td>
<td>Diacylglycerol-deacylating enzyme</td>
<td>Cd11b</td>
</tr>
<tr>
<td>Acid mucopolysaccharide</td>
<td>Fibronectin-R</td>
<td>Plasminogen activator-R</td>
<td>Cd14</td>
</tr>
<tr>
<td>α,-antitrypsin</td>
<td>G-protein subunit</td>
<td>Acetyltansferase</td>
<td>Cd16</td>
</tr>
<tr>
<td>α-mannosidase</td>
<td>Laminin-R</td>
<td>β₂-microglobulin</td>
<td>Plasminogen activator-R</td>
</tr>
<tr>
<td>Heparin-binding protein</td>
<td>Thrombospandin-R</td>
<td>Gelatinase</td>
<td>Albumin</td>
</tr>
<tr>
<td>Bactericidal permeability increasing protein</td>
<td>Plasminogen activator-R</td>
<td>Lysozyme</td>
<td>Tetranectin</td>
</tr>
<tr>
<td>β-glycerophosphatase</td>
<td>Collagenase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-glucuronidase</td>
<td>Gelatinase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cathepsins</td>
<td>Histaminase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defensins</td>
<td>Heparanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elastase</td>
<td>Lactoferrin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lysozyme</td>
<td>Lysozyme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myeloperoxidase</td>
<td>Sialidase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteinase-3</td>
<td>β-microglobulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sialidase</td>
<td>TNF-R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TNF, tumour necrosis factor.
mast cells (Peretz et al. 1994). There are four known inherited abnormalities of eosinophils. An absence of eosinophil peroxidase, an autosomal recessive trait, usually results in no clinically detectable symptoms (Wright & Gallin 1979; Pouliot et al. 1997). Chediak–Higashi syndrome manifests with large abnormal granules, seen in the eosinophil as well (Davis & Douglas 1971). A third type of abnormality was found in an individual family. It appears to be inherited in an autosomal recessive fashion. Their eosinophils demonstrated large grey inclusion bodies; however, they manifested without any clinical abnormality (Tisdale 1997). The fourth abnormality is an absence of specific granules seen in both neutrophils as well as eosinophils, presenting with repeated infections (Roos et al. 1996).

Platelets
Platelets, with their lack of a nucleus, may appear to be simple in nature but they serve a pivotal role in the regulation of haemostasis, thrombosis and inflammation. Platelet formation is accomplished by the fragmentation of megakaryocyte cytoplasm. In the circulatory system, platelets appear to be passive, smooth discs. However, they maintain the ability to recognize a site of injury, adhere to this site and serve in the activation and propagation of thrombus, as well as mediate the inflammatory pathway. Their lifespan is from 7 to 10 days and in a healthy individual their count can range from 150 000 to 440 000/μl. Platelets contain a circumferential band of microtubules, which serve to maintain the discoid shape, as well as an abundance of both actin and myosin within the platelet. These microtubules are responsible for the change in shape and spicule formation seen in platelets following activation.

Platelets are noted to have granules containing histamine, serotonin and TxA2, amongst other proteins. It is unclear whether these mediators are developed within the megakaryocyte and are transferred to the platelet via the fragmentation process or whether they are absorbed from the plasma. Regardless, these factors, when released, serve to instigate and propagate many physiological reactions associated with platelet function.

Haemostasis is the culmination of three interactive systems including vascular endothelium, blood platelets and plasma proteins of both the intrinsic and extrinsic coagulation pathways. This process serves to arrest the loss of blood from vessels that have been mechanically traumatized, for example in muscular sprains, strains and fractures. When discontinuity of a vessel occurs, a series of responses termed primary haemostasis ensues. Following trauma, the vessel wall quickly retracts and platelets immediately adhere to the subendothelial collagen. Adherence to the vessel wall prompts platelet activation, which leads to propagation of the thrombus. This is continued until occlusion of the traumatized vessel occurs. The initial adherence is mediated by von Willebrand factor found in the plasma as well as von Willebrand factor released from activated platelet and endothelial cells. Following adhesion and activation (i.e. granule release), P-selectin, a platelet granule membrane glycoprotein, translocates to the cellular surface (Berman et al. 1986; McEver 1991; Frenette & Wagner 1997). This glycoprotein mediates the adhesion of leucocytes such as monocytes and neutrophils. Activation of the platelet eicosanoid pathway occurs, leading to the formation of arachidonic acid (Serhan et al. 1996; Sarraf et al. 1997). Arachidonic acid is released where it is immediately converted to PGH2. This is then converted to TxA2, a potent vasoconstrictor (O’Rourke et al. 1997). The activated platelet undergoes a change in shape with the formation of spicules. This change allows more effective binding between platelets as well as increased binding to factor X and activation of factor VII of the extrinsic coagulation cascade.

The platelet not only serves an important role in the regulation of the coagulation cascade, but serves as an important source of vasoactive mediators as well. Following vascular injury, activation of the coagulation cascade or exposure to the basement membrane stimulates platelets to release a variety of factors. These factors include serotonin, TxA2 and histamine. Serotonin and histamine are generally released from cytoplasmic...
granules and serve as an immediate stimulant to increase vascular permeability. Thromboxane A₂, an arachidonic acid derivative formed by the breakdown of membrane phospholipids, serves as a stimulant to secondary clot formation as well as the aforementioned function of smooth muscle vasoconstriction. Interestingly, the absence of platelets appears to induce an increased state of vascular permeability. The mechanism of this is unclear, but in patients who are thrombocytopenic, spontaneous cutaneous as well mucosal bleeds are frequent.

**T-lymphocytes**

T-cells are vitally important in the normal functioning of the human immune system. In the human body, T-cells develop in the thymus. Utilizing the CD4 and CD8 receptor molecules expressed on the surface of T-cells, they can be divided into four subsets (Sprent & Webb 1987; Fink & Bevan 1995). These include CD4+/CD8+, CD4+/CD8−, CD4−/CD8+ and CD4−/CD8−. Both CD4+ and CD8-negative T-cells make up about 2% of the total thymocytes. Their function remains unknown; however, they do seem to serve as precursors to the remaining subsets. T-cells are unique—they must display maximal reactivity to an infinite number of antigens but remain complacent in the face of self-antigens. In order to facilitate this, intrathymic precursors undergo a complex process of both positive and negative selection based on their reactivity to self-peptides bound to major histocompatibility complex (MHC) molecules.

Positive selection occurs in the cortical region of the thymus. Since MHC complexes are highly polymorphic, each individual must create a population of T-cells that are capable of recognizing these molecules and reacting to them if they are bound to an antigen (Sprent & Webb 1995). The double positive cells are exposed to self–MHC complexes and those precursors to circulating T-cells that display the ability to recognize these complexes are retained. The cells that are unable to recognize these cells are allowed to die in situ (Strang et al. 1988). The cells that demonstrate an overly aggressive nature to MHC molecules are removed via a process termed negative selection, leaving only the physiologically reactive T-cell progeny to survive in the post-thymic environment. In contrast to the cortical location of positive selection, it is unclear where the process of negative selection occurs. There is evidence that supports both a cortical and medullary localization of these events (Snijdewint 1993; Neiman 1997). T-cell importance in the acute inflammatory phase has not been well documented although its absence has been shown to decrease the strength of healing collagen, while prolonged activation has been implicated in the formation of excessive scar tissue.

**B-lymphocytes**

B-cell lymphocytes are derived from the bone marrow in humans. They are a key element in immune responses to foreign antigens. There are two types of B-cell immune responses to antigens: T-cell independent- and T-cell-dependent responses. T-cell-independent responses are accomplished by the binding of an antigen, leading to cross-linking of the B-cell receptors. This stimulates proliferation and differentiation of B-cells into antibody-secreting plasma cells (Nossal 1994; Doody et al. 1996). The majority of immune responses to antigens, however, are T-cell dependent. In this process, the B-cells present antigens to T-cells in order to beseech their assistance. After being presented an antigen, the B-cell will either pinocytose or endocytose it via receptors on the cellular membrane. In the B-cell, the antigen is processed and broken down into peptides, which are then presented by the cell’s MHC molecules to antigen-specific T-cells. Following recognition of the peptide–MHC complex by the T-cell, a complex interaction occurs between the T-cell and B-cell. This interaction requires cell-to-cell contact and involves multiple receptors and ligands on both cells (Oliver & Essner 1975; Mitchell et al. 1995). Signals within this interaction are critical in the development of further immune responses, such as immunoglobulin isotype class switching and the generation of memory B-cells.
In the athlete, the amount of B-lymphocytes as well as T-lymphocytes appears to be decreased during and immediately following vigorous exercise. Some researchers have speculated that an increased incidence of upper respiratory tract infections seen in high-level endurance athletes is a result of this decrease in lymphocytes (Neiman 1997). The decrease in numbers is most probably related to the secretion of exercise-induced hormones such as cortisol (Neiman 1997). Although B-lymphocytes have been implicated in such phenomenon as exercise-induced asthma and upper airway disease in the athlete, their role in the acute phase of injury remains elusive. Considering the complex interactions seen between both plasma and cellular mechanisms of inflammation, one may speculate that B-lymphocytes serve an active role in the generation and/or regulation of the inflammatory response. Further studies are needed, however, to define the exact nature of their involvement.

Natural killer cells

Natural killer (NK) cells represent a discrete subset of the lymphoid population, differing from T-cells and B-cells in that they do not express or rearrange known receptors for antigens. NK cells can account for up to 20% of circulating lymphocytes. NK cells appear to provide a first line of defence against tumour cells and viral infections (Tracey & Smith 1978). They are characterized by the expression of two distinct membrane proteins, the CD56 receptor and the CD16 receptor. CD16 is a low affinity receptor for the Fc portion of immunoglobulin G, whereas CD56, which is analogous to the neural cell adhesion molecule (NCAM) (Tracey & Smith 1978). These cells were found to propagate cytolytic cell destruction in the absence of deliberate immunization and in the absence of MHC complexes. The recognition of these foreign cells seems to be mediated by the absence of MHC class I molecule (Ljunggren & Karre 1985; Sprent 1993). The presence of such molecules appears to serve as an inhibitory stimulus to the NK cells. Activation of NK cells not only results in their cytolytic activity, but also stimulates the release of cytokines that can exert a regulatory role in the immune response and inflammation (Trinchieri 1989; Bellone et al. 1993).

The complement system

The complement system is partially responsible for the recognition and destruction of pathogens and altered host cells. In this respect it is a highly complex system that plays an imperative role in both the innate immune system as well as a primed immune system ready to react against a pathological insult. The complement system possesses the ability to directly recognize and eliminate pathogens or damaged host cells. The mechanisms by which this is accomplished will be considered below.

The complement system is made up of more than 30 plasma and cell membrane proteins. When activated by a pathogen or injury, a precisely regulated series of interactions, not only within the complement proteins but also with the pathogen and cell membranes, initiates a series of reactions that can be divided into three stages. The first is the recognition and initiation of one of the two complement-activating pathways (classic pathway and alternate pathway.) The second is C3 binding and amplification; the third and final phase is the membrane attack pathway. The process of activation of the classic pathway involves recognition of the inflammatory agent by the first component of complement (C1). The C1 component consists of three distinct proteins: C1q, C1r and C1s. The C1q protein attaches itself to the Fc component of the instigating immunoglobulin. C1r and C1s are activated with this binding and in turn react with C4 and C2 forming two products, an anaphylatoxin C4a and C3 convertase (C4b and C2a). Cs3 convertase cleaves the C3 molecule forming C3a and C3b (a second anaphylatoxin). The resulting complex binds to the C5 molecule initiating hydrolysis of this molecule forming C5a and C5b. C5b will attach itself to target cell membranes and acts to facilitate binding of C6, C7 and C8 as well as initiating the polymerization of multiple C9 molecules. This forms a macromolecule known as the membrane
attack complex (Fig. 2.4). Following the assembly of the membrane attack complex on the cell membrane of the target cell, small cylindrical holes are formed within the membrane leading to cell lysis (Cooper 1985; Muller-Eberhard 1986).

Cell products, bacteria or foreign material may also activate the alternative pathway. In this pathway, C3 binds with two plasma proteins, factor B and factor D, with a resultant product Bb catalysing the conversion of C3 to C3a and C3b. With the presence of Bb, C3b functions in an amplification reaction stimulating C3 convertase to form more C3a and C3b. C5 convertase is formed with the generation of C5a and C5b. The membrane attack complex is then formed as in the classic pathway.

When the classic pathway is utilized, the formation of anaphylatoxins (C3a, C4a and C5a) is an important step. Each of these molecules is capable of causing smooth muscle contraction as well as increasing vascular permeability. C3a and C5a also possess the ability to activate mast cells and basophils leading to their degranulation, with the release of histamine, PAF and serotonin. These mediators serve to further increase vascular permeability. The anaphylatoxins appear to mediate the smooth muscle contraction seen in bronchial spasm. This is accomplished by one of two pathways, the degranulation of the mast cell or a second pathway utilizing an arachidonic acid derivative. C5a also serves an important role as a chemotactic agent for leucocytes. With the initiation of the complement system, local production of C5a serves to recruit neutrophils and monocytes to sites of tissue injury. This cleavage component also has the ability to stimulate neutrophil degranulation and superoxide anion production (Jose et al. 1981).

Mechanical trauma to tissues such as fractures, muscular strains or sprains leads to the exposure of the basement membranes, which in turn activates factor XII (Hageman factor). This activated Hageman factor not only cleaves plasminogen and prekallikrein, generating plasmin and kallikrein, respectively, but also has the capability of activating the alternative complement pathway. Plasmin maintains the capability to induce increased vascular permeability as well as utilizing the anaphylatoxins C3a and C5a to induce changes in microvascular permeability. In this way, acute traumatic injuries utilize the complement system to mediate acute changes in vascular permeability as well as to initiate the inflammatory response.

The complement system is regulated by several mechanisms including: (i) the spontaneous decay of active complexes or cleavage products; (ii) the inactivation of specific components by proteolysis; and (iii) the binding of active components by plasma proteins. Factor I, a plasma protein, is responsible for the down-regulation of both C3b and C4b (Liszewski & Atkinson 1993). A second plasma protein, C1 esterase inhibitor, serves to regulate the activation of the classic pathway by binding to both C1r and C1s making them inactive (Liszewski & Atkinson 1993). Serum carboxylase N can inactivate the anaphylatoxins C4a, C3a and C5a. Factor H is another plasma protein, which binds to the C3b protein, making it more susceptible to cleavage by factor I.
The importance of the complement system can be seen by examining patients with congenital deficiencies. The absence of any of the classic complement pathway (CCP), components results in systemic lupus erythematosus (SLE) (Liszewski et al. 1989). In addition to SLE, the absence of C3, factor H and factor I appear to make an individual more susceptible to repeated pyogenic infections by strains of *Staphylococcus*, *Streptococcus*, pneumococcus and other organisms (Morgan & Walport 1991). Whereas impaired CCP activation leads to autoimmune disease, immune complex disease and repeated infections, uncontrolled CCP activation occurs in the absence of regulatory proteins. It appears that an unregulated CCP is involved in the pathogenesis of hereditary angioedema (Table 2.4) (Storkus et al. 1987; Davis 1988).

**Table 2.4 Patterns of disease in complement-deficient patients.**

<table>
<thead>
<tr>
<th>Deficient component</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1q, C1r, C1s, C4, C2</td>
<td>SLE, autoimmune diseases and recurrent pyogenic infections</td>
</tr>
<tr>
<td>C3, factor H, factor I</td>
<td>Recurrent pyogenic infections and immune complex diseases</td>
</tr>
<tr>
<td>C5, C6, C7, C8</td>
<td>Recurrent neisserial infections</td>
</tr>
<tr>
<td>C1 inhibitor</td>
<td>Hereditary angioedema</td>
</tr>
<tr>
<td>CR3, CR4</td>
<td>Severe immunodeficiency, leucocyte dysfunction and recurrent infections</td>
</tr>
</tbody>
</table>

The importance of the complement system can be seen by examining patients with congenital deficiencies. The absence of any of the classic complement pathway (CCP), components results in systemic lupus erythematosus (SLE) (Liszewski et al. 1989). In addition to SLE, the absence of C3, factor H and factor I appear to make an individual more susceptible to repeated pyogenic infections by strains of *Staphylococcus*, *Streptococcus*, pneumococcus and other organisms (Morgan & Walport 1991). Whereas impaired CCP activation leads to autoimmune disease, immune complex disease and repeated infections, uncontrolled CCP activation occurs in the absence of regulatory proteins. It appears that an unregulated CCP is involved in the pathogenesis of hereditary angioedema (Table 2.4) (Storkus et al. 1987; Davis 1988).

**Response to injury**

Numerous changes occur within the human body after the onset of an inflammatory state. These changes are not only localized to the site of the pathology but also involve numerous organ systems. The acute phase response represents a state in which the body modifies its normal internal environment to appropriately respond to an inciting pathology. Cytokine-induced changes in plasma protein synthesis, the neuroendocrine and haematological systems, metabolic processes, and non-protein plasma components mediate these local as well as distant changes.

The initial response to tissue injury occurs primarily in the vasculature, namely the capillaries and postcapillary venules. Normal vascular anatomy consists of a layer of endothelium connected by tight junctions and a basement membrane composed of type IV collagen, glycosaminoglycans and glycoproteins. This vascular anatomy serves to preserve the normal relationship between the tissue and the circulating plasma and cellular components, which is one of mutual exclusion. After tissue injury, an initial rapid step is activation of the endothelial cells. The endothelium, when activated, has the ability to change its surface properties and become adhesive for both platelets and leucocytes. Shortly after acute trauma, the endothelium expresses an adhesion protein, P-selectin. This protein binds both polymorphonuclear leucocytes and monocytes; it is this binding or tethering that accounts for the rolling of leucocytes along the endothelium. The activated endothelium also stimulates the production of PAF. The synthesized PAF is directed towards the surface of the endothelial cell, but is not released. This localization of PAF to the surface induces tight binding of the leucocytes with their subsequent emigration, and also primes them for degranulation. PAF has also been implicated in the activation of passing platelets by the bound leucocytes. The platelets, once activated, release mediators such as TxA₂, serotonin and histamine, leading to the formation of gap junctions in the endothelium and thus facilitating leucocyte migration (Fig. 2.5). The formation of gap junctions leads to a disruption of endothelial continuity and exposing of the basement membrane, which in turn further activates platelets. This leads to a self-propagating cycle which is vital in the repair of damaged tissue.
The response of the vasculature to injury is not only one of increased permeability, but also one of dynamic vasodilatation and vasoconstriction. At sites of injury, vasoactive mediators, both cellular and plasma derived, bind to specific receptors located on the endothelial cells as well as on the vascular smooth muscle, propagating either vasoconstriction or dilatation. The effects of vasodilatation and constriction have variable effects on the tissue, depending on their site of action. Vasodilatation of the arterioles leads to increased blood flow to a tissue, i.e. increased oedema, whereas constriction leads to decreased blood flow, i.e. decreased oedema. Conversely, dilatation of venules decreases capillary hydrostatic pressure leading to decreased oedema formation, whereas constriction leads to increased hydrostatic pressure in the capillaries and subsequent increased tissue oedema.

The postcapillary venule appears to be the primary site of endothelial changes for many of the vasoactive mediators. As mentioned above, during the time immediately following an injury, a complex cascade of biochemical events is initiated, leading to cell contraction, loss of tight junctions and formation of endothelial gaps. This series of events leads to increased vascular permeability with the subsequent extravasation of intravascular fluids into the injured tissue. This process is dynamic and reversible with maximal permeability occurring within 10–30 min after injury. Normal vascular anatomy is regained usually within an hour.

Several modalities have been demonstrated to alter the course of inflammation, both in the acute and chronic setting. Non-steroidal anti-inflammatory drugs (NSAIDs) serve as membrane stabilizers, thus dampening the effects of the arachidonic acid cascade. The use of NSAIDs has not been proven to speed wound healing and in fact has proved to be of minimal benefit in the treatment of acute injuries. Corticosteroids act early in the inflammatory cascade by blocking the release of arachidonic acid. There have been no clinical studies that demonstrate the oral use of corticosteroids in acute sports trauma to be beneficial in the recovery of an athlete. Corticosteroid injection therapy, on the other hand, has been implied in animal models to be advantageous to wound healing, but no conclusive data have been obtained. Injection of corticosteroids should be limited in acute macrotrauma secondary to its catabolic effect, and injections around tendinous and ligamentous structures must be weighed carefully secondary to the deleterious effects on those structures. Physical modalities such as thermotherapy and cryotherapy have been used in the treatment of acute inflammation. Thermotherapy, or heat application, has been shown to relieve the secondary symptoms of inflammation such as muscle spasm. Cryotherapy appears to reduce the local tissue temperature and local blood flow. This serves to decrease the subsequent

![Diagram of vasculature response to injury](image)

---

**Fig. 2.5** Following tissue injury, exposed vascular collagen stimulates platelet aggregation and activation (a). Activation stimulates the release of mediators leading to gap junction formation as well as leucocyte aggregation to the site of injury, with subsequent migration into the tissues (b).
oedema and may alter secondary necrotic effects
of the initial trauma.

Acute phase proteins are defined as plasma
proteins whose amount of production is either
increased or decreased by a factor of 25% in the
face of an inflammatory stimulus (Izumi et al.
1994). Some of the well-known human plasma
proteins include ceruloplasmin, the complement
components C3 and C4, C-reactive protein (CRP)
and serum amyloid (Kushner 1982; McCarty 1982).
These proteins are termed positive acute phase
proteins in that their production is markedly
increased. CRP and serum amyloid is usually
present in plasma in trace amounts. After an
inflammatory insult their concentration may
increase 1000-fold (Fig. 2.6) (McCarty 1982). The
synthesis of negative acute phase proteins is
by definition decreased by at least 25%. Long
recognized negative acute phase proteins include
transferrin and albumin. The functions of these
acute phase proteins are as varied as the proteins
themselves. Some, such as complement pro-
teins, are essential in the elimination of foreign
pathogens; the functions of other proteins, such
as CRP, remain elusive, but do provide us with
an adequate marker of the inflammatory status.

Neuroendocrine changes seen in the acute phase
response include fever, somnolence, anorexia and
altered synthesis of endocrine hormones includ-
ing corticotrophin-releasing hormone (CRH),
glucagon, insulin, adrenal catecholamines, growth
hormone, adrenocorticotropic hormone (ACTH),
thyroid-stimulating hormone, aldosterone and
vasopressin (Chrousos 1995; Boelen et al. 1997).
For the most part, these changes in the neuro-
endocrine system do not appear to be functional,
rather unwanted collateral effects of the medi-
ators of inflammation. Fever has been discussed
in the previous section on prostaglandins, but in
addition to the effect of PGE on the stimulation
of fever, it appears that the cytokine, IL-6, stimu-
lates the induction of fever via its effects on the
thermoregulatory centre (Dinarello 1997). The
production of IL-6 occurs in the endothelium of
hypothalamic organs and is a response to IL-1α
and TNF-α, common proinflammatory mediators
(Lucino & Wong 1996).

Not only the inflammatory process, but also the
stress experienced by the body, affects the neuro-
endocrine system. The endocrine abnormalities
experienced by the body are the result of a com-
plex interaction between inflammation-associated
cytokines—predominantly the hypothalamic–
pituitary–adrenal axis (Chrousos 1995). Common
hormone increases are seen in insulin, glucagons,
cortisol and ACTH, as well as others (Patel &
Neuberger 1993). IL-1, IL-6 and TNF-α appear to be amongst the most potent stimulators of the hypothalamic–pituitary–adrenal axis by stimulating the production of CRH and arginine vasopressin, with a consequent increased production of ACTH and cortisol. The sympathetic and adrenomedullary systems respond to both pro-inflammatory mediators as well as to endocrine changes, with the secretion of neurotransmitters. These further lead to alterations in the neuroendocrine environment associated with the acute phase response.

Lethargy, anorexia and somnolence appear to be common findings during the acute phase response. In experimental models IL-1 and TNF-α both induce a somnogenic effect on rabbits when injected into their cerebral ventricles (Surh & Sprent 1994). IL-1 has been demonstrated within the central nervous system of humans and appears to play a central role in the somnogenic response in them as well (Surh & Sprent 1994; Leon et al. 1997). Several cytokines are implicated in the pathogenesis of anorexia. In animal models, IL-1-induced secretion of IL-6 appears to stimulate an anorexic state; however, clinical trials to support these findings are lacking in humans (Zeidler et al. 1992; Fattori et al. 1994; Leon et al. 1996). Other animal studies have implicated local production of TNF-α as well as IL-1α-induced secretion of leptin (an acute phase reactant) as a stimulant to anorexia (Zeidler et al. 1992; Ryan 1997). In truth, as with most types of inflammatory reactions, its causes are most likely multifactorial and remain unclear.

Anaemia of chronic inflammation appears to be secondary to a decreased production of red blood cell progenitors, as well as a decrease in erythropoietin synthesis (Means 1995). Both of these appear to be mediated somewhat by inflammatory cytokines. Some of these cytokines decrease the response of the red blood cell progenitors to erythropoietin, whereas others decrease the synthesis of erythropoietin (Faquin et al. 1992). The major component may be the decreased production of erythropoietin, since studies have shown that anaemia secondary to chronic inflammation can be overcome with the administration of erythropoietin alone (Zabucchi et al. 1992).

Metabolic changes seen in the face of chronic inflammation include the loss of body mass and altered lipid metabolism. As with neuroendocrine changes, these are probably an untoward effect of circulating cytokines rather than a desired response. Decreases in the amount of skeletal muscle, fat tissue and bone mass result from the effects of numerous cytokines including IL-1, IL-6, TNF-α and interferon gamma (IFN-γ) (Espm et al. 1995; Takahashi et al. 1996). IL-1 and IL-6 appear to be the primary cytokines responsible for the loss of skeletal muscle, which is a result of decreased protein synthesis as well as increased muscle proteolysis (Cannon 1995). Alterations in lipid metabolism results in the loss of fat tissue and decreases in circulating high-density lipoproteins, as well as increases in serum triglycerides, very low-density lipoproteins and low-density lipoproteins (Liao & Fioren 1993; Feingold et al. 1994; Banka et al. 1995). Unlike the cytokines responsible for muscle degradation, the cytokines responsible for the altered lipid metabolism remain elusive. Studies have demonstrated that chronic injections of IL-6 in Rhesus monkeys appear to induce a state of hypocholesterolaemia, but the correlation to human lipid metabolism remains unclear. Immunosuppression has also been implicated as an acute phase phenomenon (Ettinger et al. 1995).

The onset of an inflammatory process is accompanied by numerous physiological reactions at sites distant from the initial insult. The acute phase response to injury represents a complex interaction of numerous organ systems. These interactions appear to be mediated by inflammation-associated cytokines and influenced by modulators of cytokine function as well as endocrine hormones and other circulating factors. Although these changes are frequently seen in association with one another, it is common to have variable responses to inflammation on an individual basis.

The process of tissue repair

When faced with inflammation, the body is forced to initiate the reparative process. The re-establishment of physiologically functional tissue is imperative for the return to activity levels
similar to the preinjured state. A coordinated process of cell migration and proliferation directed by specific biochemical mediators facilitates this repair of damaged tissue. This process utilizes various aspects of the humoral and cellular defence mechanisms to alleviate the inciting pathology.

Following an injury in which the vasculature is disrupted, the process of tissue repair begins immediately. As mentioned above, platelet adherence to the exposed collagen stimulates activation with the subsequent release of mediators such as serotonin, histamine and TxA₂, among others. Once the acute haemorrhage is controlled by the activated platelet and clotting pathways, the migration of inflammatory cells into the region of damage begins.

In the immediate postinjury period, neutrophils predominate the immigrating leucocytes. They can be detected within 1 h after injury and peak in approximately 24–48 h. Recruitment of these cells is mediated by the process of rolling (via selectins), adhesion (via integrin) and migration through the endothelium (Menger & Vollmar 1996). Immigration into the site of injury is stimulated by complement components (C5) as well as factors released by the activated platelets (Marder et al. 1985).

The function of these neutrophils is to clear the wound of fibrin as well as the initiation of inflammation via the release of proinflammatory cytokines (Grinnell & Zhu 1994). As mentioned earlier, the regulation of cytokines is primarily at the transcription level. Messenger RNA of TNF-α is apparent at the site of injury within 12 h, peaking at 72 h. This increase in mRNA may last for up to 5 days after the injury (Feiken et al. 1995).

As the number of neutrophils begins to decrease, a concomitant rise in the number of macrophages is seen. This increase in macrophages seems to peak in 5–7 days. The recruitment of these circulating macrophages to the site of injury is secondary to factors released by both platelets and neutrophils (Cromack et al. 1990; Ham et al. 1991). These macrophages serve as a major source of growth factors and cytokines that recruit and activate additional macrophages and fibroblasts. Macrophages function in debridement of the pathology, recruitment of other inflammatory cells, cell proliferation, and regrowth of peripheral nerves (Leibovitch & Ross 1975; Platt et al. 1996; Fritsch et al. 1997).

T-lymphocytes have been shown to be present at sites of injury from the first hour post-injury and peaking between 7 and 14 days later (Fig. 2.7) (Martin & Muir 1990). Their function is not clearly understood, although their absence has been shown to lead to decreased wound strength and lower amounts of collagen deposition. Conversely, prolonged activation of T-cells appears to lead to excessive fibrosis as seen in keloid formation (Borgognoni et al. 1995). This can be logically expanded to areas of chronic inflammation, where increased amounts of fibrosis are seen, presumably secondary to prolonged T-cell activation (Peters et al. 1986).

Following the recruitment of leucocytes and fibroblasts, a tissue termed granulation tissue is formed. This tissue was initially termed granulation tissue secondary to the granular appearance of the newly forming blood vessels. Granulation tissue develops from connective tissue surrounding the site of injury as well as the recently

Fig. 2.7 Following an injury to biological tissue, migration of characteristic leucocytes into the damaged tissue follows a consistent pattern. The length of their presence depends on the amount of trauma.
immigrated leucocytes, fibroblasts and myofibroblasts. Over a period of time, this granulation tissue is gradually replaced with more organized collagen, which remodels to eventually form organized scar tissue.

**Conclusions**

Although inflammation is often thought of as an unwanted effect of trauma or overtraining, it represents a physiological state of repair that’s purpose is to return the individual to a state of functional recovery. The development of an inflammatory reaction is multifaceted, utilizing many of the body’s systems to mediate its purpose. Complex interactions between cellular and plasma-derived components serve to initiate, propagate and regulate the inflammatory response.

Shortly after a physiological insult in which disruption of the vasculature occurs, platelets aggregate to the site of injury. Via interactions between the platelets, exposed collagen and endothelial cells, factors are secreted which serve to propagate numerous functions. Vasoconstriction at the site of injury initially serves to decrease the blood flow, thus minimizing the amount of blood loss at the site of injury. Gap junctions are formed within the endothelium to facilitate the transendothelial migration of circulating leucocytes in order to gain access to the site of pathology. The local secretion of cytokines and chemokines, which serve as chemottractants, assists the accumulation of leucocytes at the site of injury. Depending on the type of pathological insult, the type and amounts of cytokines, as well as other stimulators of cellular migration (complement proteins, antigen–antibody complexes), vary. It is the variability of these signals, which dictate the type of cellular components recruited to the site of injury.

Regardless of the cellular mediators, the ultimate goal is the same—removal of diseased tissue with the subsequent regeneration of tissue, functionally equivalent to the prediseased state. Depending on the type and length of the insult, variable amounts of newly generated tissue will be deposited. This newly formed tissue usually takes the form of collagen with the physiological and histological appearance of fibrosis. If the pathological insult is small, the ratio of fibrosis to normal tissue remains small and the functionality of the original tissue is left intact. In instances of prolonged pathological insult, the ratio of fibrosis to normal tissue is elevated, thus compromising the host tissue function. It is for this reason that the athlete as well as the physician attempt to minimize the amount of inflammation.

**References**


Chapter 3

Skeletal Muscle Regeneration After Injury: Cellular and Molecular Events

ANDRÉ-XAVIER BIGARD AND EMMANUELLE FINK

Introduction

Muscle damage is frequently observed after sports injuries. The ability of skeletal muscle to regenerate after injury is one of its major characteristics and myofibres are repeatedly damaged and repaired during adult life. At the anatomical level, injuries are commonly divided into shearing injuries, in which both myofibres and framework are torn, and injuries in situ, in which only myofibres are damaged, such as after repeated eccentric contractions. Shearing injuries result mainly from direct trauma to skeletal muscle or strain injuries, while in situ injuries follow exhaustive exercise, the application of local anaesthetics or are caused by diseases. The treatment and prognosis of muscle injuries vary widely according to the severity and extent of the trauma, but regardless of the origin of the injury, skeletal muscle will regenerate.

In response to injury, skeletal muscle regenerates following two types of regeneration. Epimorphic regeneration is mainly found in amphibians and gives rise to an entire new limb after amputation (Carlson 1970). In contrast to amphibians, skeletal muscle regenerates in mammals utilizing the remnants of the original myofibre complex. Here we review and discuss some important cellular events which follow muscle damage in mammals, and epimorphic regeneration will not be treated further. Muscle regeneration has been broadly divided into two types: continuous and discontinuous regeneration (Hudgson & Field 1973). Discontinuous regeneration, of embryonic type, is characterized by the formation of newly regenerated myofibres from myoblastic cells. In contrast, continuous regeneration, of budding type, results in the outgrowth from the end of the section of a partially damaged myofibre. Continuous regeneration is less understood and will not be treated in this review.

Before describing the several phases of the healing process, it is important to summarize some aspects of the structure of satellite cells which represent stem cells providing myoblasts for regeneration. We will pay particular attention to the molecular events of activation and differentiation of satellite cells, and their clinical relevance now and in the future.

Structure of adult skeletal muscle

Muscle fibres are enveloped by the basal lamina. Between the basal lamina and the sarcolemma of the myofibres, are the satellite cells (Fig. 3.1). These unspecialized mononucleated cells, which were first described by Mauro (1961) in frog muscle, are known to have myogenic potential and to mediate the postnatal growth of skeletal muscle (Rosenblatt et al. 1994; Schultz 1996). In addition to their role in muscle hypertrophy and postnatal growth, there is now a large body of evidence indicating that satellite cells function as stem cells that provide myoblasts for muscle regeneration in adults. The exact origin of satellite
cells has not been clearly identified. However, it seems that these myogenic cells stem from early myoblasts that were not incorporated in the developing syncitia, and remained in the surface of all fibres (Schultz 1989). Satellite cells adjacent to mature myofibres do not express specific markers of committed myoblasts such as early-acting myogenic regulatory factors (MRFs), Myf-5 and MyoD, growth factors, or other known markers of terminal differentiation. This finding is consistent with the hypothesis that satellite cells are stem cells with an identity distinct from that of myoblasts. The microenvironment of satellite cells, available growth factors and MRFs play a pivotal role for the generation of committed myoblasts.

There is now available evidence that satellite cells are divided into subclasses based on the fibre type in which they lie. Moreover, there is also evidence that satellite cells form a heterogeneous population based on their profile of gene expression (Cornelison & Wold 1997). It has been demonstrated that satellite cells have intrinsic properties dependent on their muscle origin, and independent of environmental cues (Dolenc et al. 1994; Martelly et al. 2000). Satellite cells from either slow or fast muscles present significant differences on the expression of several protein families such as metabolic enzymes, hormonal receptors or capacity to express MRFs. These differences could have consequences on the capacity of slow and fast muscles to regenerate after injury.

Satellite cells are normally in a non-proliferative, quiescent state, but are activated in response to muscle injury (Schultz et al. 1985; Grounds 1998). Multiple rounds of proliferation of these stem cells occur after their activation, giving rise to myogenic precursor cells (Grounds & Yablonka-Reuveni 1993). The number of quiescent satellite cells is dependent on age and muscle fibre type. Skeletal muscle in young animals contains a higher concentration of satellite cells than in adults (Schultz 1989).
The percentage of satellite cells within muscles decreases with age, and more rapidly from birth through sexual maturity than in adults. The decrease in the percentage of satellite cells with ageing is the result of an increase in myonuclei in both oxidative and glycolytic myofibres, and a decrease in total number of satellite cells. On the other hand, the satellite cell density is higher in oxidative than in glycolytic tissue, at both whole muscle and single fibre levels (Gibson & Schultz 1982, 1983). This heterogeneity in satellite cell content between muscle fibre types has been related to an increase in satellite cell density with the proximity of capillaries and motor neurone junctions. While the population of satellite cells decreases with age, their number remains relatively constant over repeated cycles of degeneration–regeneration. This finding highlights the inherent capacity of these stem cells for self-renewal (Gibson & Schultz 1983). At least three models have been suggested to account for the self-renewal of satellite cells, involving either early-acting MRFs, an asymmetrical satellite cell division and/or the de-differentiation of committed myogenic precursor cells (Seale & Rudnicki 2000).

Skeletal muscle injury is generally followed by a series of processes included in three phases: a degenerative phase, a regenerative phase and a maturation phase (Fig. 3.2) (Plaghki 1985). Many experimental results suggest that basic mechanisms underlying the cellular responses to acute trauma are well conserved regardless of the type of initial injury. Although the sequence of cellular responses of injured muscle regeneration is roughly constant and well determined, the time-course of muscle regeneration processes are tightly related to the types of muscle injury (Carlson 1973). Using drastic experimental models, it has been shown that the regenerative process takes place earlier after in situ injuries without destruction of the basement membrane, than after denervation/devascularization of muscle (Lefaucheur & Sebille 1995).
The degenerative phase

First phase of the degenerative process

This is also known as non-inflammatory or intrinsic degeneration, and follows the initial damage of the muscle fibres. An extensive disruption of the structural components of the muscle is first observed at the site of injury. The extent of anatomical damage of the muscle tissue varies significantly according to the nature of the injury and its severity. With regard to exercise-induced muscle injury, many morphological studies showed that the Z-discs are the most vulnerable structures (Fridén & Lieber 2001). Other intracellular anomalies have been reported according to the severity of the injuring process, and damage of the sarcolemma, T-tubules and the cytoskeletal system are commonly observed after exercise-induced muscle injuries. More severe trauma, including strains, contusions, prolonged ischaemia and muscle lacerations are associated with the crush and tear of myofibres, and often the death of nearly all myonuclei (Fig. 3.3).

The early responses to initial damage consist of catabolic events that result in autolysis of the damaged muscular components. A loss of intracellular calcium homeostasis has been shown to activate calcium-dependent proteases referred to as calpains (Armstrong et al. 1991). Such enzymes are able to cleave myosin, α-actinin, vinculin and many other contractile filament components and metabolic proteins in the muscle cell. Calpain activation appears thus to be a key occurrence of damaged muscle autolysis. Most of these alterations become prominent 4–8 h after initial injury. Satellite cells, which are dormant under the basement membrane of myofibres, withstand the initial damage and its early consequences, and then enter into a cycle of activation–proliferation.

Second phase of muscle degeneration

This phase is also called extrinsic or inflammatory degeneration, and involves the removal of all traces of the originally damaged myofibres. This phase of muscle repair includes phagocytosis, which is secondary to an inflammatory cell response, and is dependent on the presence of invading macrophages (Fig. 3.2).
Neutrophils and mononucleated cells

Neutrophils and mononucleated cells accumulate at muscle injury sites. Neutrophils respond early to muscle injury. The neutrophil population increases significantly within 1–6 h and declines 9–12 h after damage (Orimo et al. 1991). In certain cases this rapid increase in the neutrophil population endures for a few days (Fielding et al. 1993). This early increase in the number of inflammatory cells may result either from chemotaxis of specific cells to the site of muscle damage and/or from mitogenesis of inflammatory cells initially present in skeletal muscle near the myofibres. There is experimental evidence that chemotaxis is the primary mechanism by which the number of inflammatory cells increases in injured muscle (Bintliff & Walker 1960).

The exact nature of the substances acting as chemoattractants for inflammatory cells remains a matter of debate (Tidball 1995). Several molecules released from injured muscles have been viewed as active substances for leading inflammatory cells from the circulatory system to the injury site (Table 3.1) (Chen & Quinn 1992). Basic fibroblast growth factor (bFGF or FGF-2) may be released through membrane lesions, but while it is capable of attracting myogenic cells, there is no experimental evidence to support a role in attracting inflammatory cells after injury. It is thus unlikely that bFGF is a key chemoattractive stimulus for inflammatory cells. In contrast, platelet-derived growth factor (PDGF), released from injured muscle, is known as a mitogen and chemoattractant for inflammatory cells. However, its low concentration in myofibres and short half-life make uncertain its role as a chemoattractant to mediate the inflammatory response to muscle injury (Bowen-Pope et al. 1984). Myostatin, a new muscle-specific growth-inhibiting factor, has been recently suggested to act as a chemoattractant for phagocytes and inflammatory cells (Kirk et al. 2000). The specific role played by many other soluble factors such as cytokines as chemoattractants in damaged muscle remains unknown to date. Additional data suggest that the early activation of resident macrophages or fibroblasts could initiate the release of other chemoattractant substances, leading to the translocation of inflammatory cells within injured muscle. Once resident macrophages and/or fibroblasts are activated by chemoattractant substances, these later inflammatory cells could initiate chemotaxis of additional cells by a wound hormone unidentified to date.

The inflammatory response to muscle injury is initiated by the local proteolysis of extracellular matrix molecules. Fibroblasts probably play a significant role in degrading extracellular matrix molecules, mainly by an increase in collagenase expression. PDGF and cytokines

Table 3.1 The main effects of several growth factors and cytokines on chemotaxis, proliferation and differentiation of satellite cells.

<table>
<thead>
<tr>
<th>Growth factor or cytokine</th>
<th>Basic effects on muscle</th>
<th>Chemotaxis</th>
<th>Proliferation</th>
<th>Differentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>bFGF</td>
<td>+/−</td>
<td>++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>PDGF</td>
<td>+</td>
<td>++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>IL-1</td>
<td>+ (lymphocytes)</td>
<td>++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>IL-6</td>
<td></td>
<td>++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>LIF</td>
<td></td>
<td>++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>IGF-I, IGF-II</td>
<td></td>
<td>++</td>
<td>++</td>
<td>−</td>
</tr>
<tr>
<td>TGF-β</td>
<td>+</td>
<td>+/−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>TNF-α</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>NGF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
detected in injured muscle, such as interleukin-1 (IL-1), may stimulate collagenase secretion by fibroblasts and thus contribute to the proteolysis of extracellular matrix molecules (Dayer et al. 1986).

**Polymorphonuclear lymphocytes and macrophages**

Polymorphonuclear lymphocytes and macrophages migrate to sites of tissue damage within a few hours of the injury (Fig. 3.4). Macrophages constitute the major type of invading cell during the inflammatory response to muscle injury. They are the predominant inflammatory cell type 1 day postinjury, and more than 50% of inflammatory cells are macrophages following exercise-induced injury (Round et al. 1987). Many investigations have confirmed the marked increase in macrophages within damaged muscle and emphasized the major importance of these cells in muscle inflammation. Distinct subclasses of macrophages are involved in the development of muscle inflammation. Macrophages normally present in the circulatory system increase significantly early after injury, while macrophages normally resident in muscle tissue increase thereafter (St Pierre & Tidball 1994). It is thus likely that macrophages increase in damaged tissue both by emigration from the circulatory system (mainly by chemotaxis) and also by cell division of resident tissue macrophages.

**Functional role of inflammatory cells in damaged muscle**

Neutrophils rapidly invade the injury site and are associated with early promotion of the inflammatory response. They are involved in the release of cytokines or soluble factors that can attract and activate additional inflammatory cells. As discussed above, fibroblasts are candidates as mediators of the chemotaxis of inflammatory cells, but above all can cause proteolysis of extracellular matrix proteins, one of the early stages of inflammation. Two pivotal roles for macrophages have been proposed based on observations that
myogenesis is markedly impaired in the absence of macrophage infiltration (Lescaudron et al. 1999).

**Phagocytic role of macrophages**

Macrophages and inflammatory cells play a prominent role in removing all traces of the damaged myofibres and consensus has been reached in considering that the phagocytic role of these inflammatory cells is their most prominent function during the recovery from muscle injury. This response of inflammatory cells is an important component of the successful repair of injured muscle. Treatment with non-steroidal anti-inflammatory drugs (NSAIDs) is known to impair the recovery of damaged muscles. It has been shown that rats treated with NSAIDs exhibit a slower muscle regeneration than animals receiving placebo (Almekinders & Gilbert 1986). The impaired recovery in muscles of treated rats is mainly attributed to a less phagocytic activity in damaged muscle, a slow down in the removal of cellular debris and a decrease in the synthesis of cytokines involved in successful muscle repair. These data clearly demonstrate that macrophages and neutrophils are directly involved in the process of muscle repair, and inhibiting the early activation and infiltration of inflammatory cells affects muscle fibre repair. As a consequence, the early treatment with NSAIDs after muscle trauma has a bad prognosis. Other situations where phagocytosis by macrophages is impaired, such as ageing, affect the success of muscle regeneration (Zacks & Sheff 1982).

Macrophage populations comprise distinct subclasses of cells. Macrophages, including the ED1 antigen, which are associated with lysosomal membranes, also called ED1$^+$ macrophages, are present at high concentrations in areas that contain necrotic fibres. They have been shown to have a high phagocytic activity and to remove cellular debris in injured tissue (Honda et al. 1990; McLennan 1993). This specialized subpopulation of macrophages declines in muscle after the phagocytic stage is completed.

**Macrophages and the secretion of growth factors and cytokines**

The success of muscle regeneration also relies on functions other than phagocytosis that may be mediated by macrophages. Although restricted in number, resident macrophages could provide growth factors and cytokines important for the early development of the inflammatory response in damaged muscle, such as transforming growth factor beta (TGF-$\beta$), IL-1, IL-6 and tumour necrosis factor alpha (TNF-$\alpha$) (Rappolee & Werb 1992). TNF-$\alpha$ is known to induce skeletal muscle protein breakdown, angiogenesis and the production of components of inflammation. IL-1 is expected to exert important influences on the development of inflammation. This cytokine is a mitogen and chemoattractant for fibroblasts, T-lymphocytes and B-lymphocytes, and increases the cytotoxicity of macrophages. IL-1 also induces the expression of other inflammatory cytokines such as IL-6 and interferons. The postexercise increase in serum and muscle IL-1 is rapid, and suggests a post-translational control of the secretion of this inflammatory cytokine (Tidball 1995). The early response of IL-1 to muscle damage appears to be mainly attributable to an increase in its production by each activated cell, rather than by a proliferation of inflammatory cells.

Taken together, these findings show that besides their role in phagocytosis, non-muscle cells such as macrophages may play pivotal and essential roles in mediating muscle repair. The first invading subpopulation of macrophages phagocytoses tissue debris, while together with neutrophils the macrophages are involved in the early inflammatory response, mainly through the secretion of soluble factors such as cytokines and local growth factors.

**Skeletal muscle regeneration**

The regeneration of myofibres begins after phagocytosis of cellular debris. The first step of regeneration includes the activation of satellite cells from their relatively dormant state, giving rise to myogenic precursor cells (Fig. 3.2). Thereafter, they
proliferate, differentiate into myoblasts, and fuse to form multinucleated myotubes that then differentiate into myofibres. Myoblasts are myogenic precursor cells expressing myogenic markers. This later step of regeneration leads to mature fibres comprising adult isoforms of the several families of proteins, consistent with the expected muscle phenotype (Plaghki 1985). The activation of satellite cells and their differentiation into myoblasts does not begin until the necrotic debris within the regenerating basal lamina cylinders have been removed by macrophages (Hurme et al. 1991). Most myoblasts giving rise to new myofibres arise from local satellite cells lying underneath the basal lamina. Although this problem has not been entirely settled, there appears to be little recruitment of satellite cells from adjacent muscles and, in most cases, repair of a muscle is the responsibility of the intrinsic satellite cell population of the damaged muscle (Schultz et al. 1986).

Activation of satellite cells

After injury, the satellite cells take several hours to be mobilized. Several non-selective markers are used to characterize the proliferating satellite cell pool, such as the incorporation of $^{3}$H-thymidine or bromodeoxyuridine. Autoradiographic studies demonstrated that satellite cells were labelled by $^{3}$H-thymidine 15–20 h after crush injury (Schultz et al. 1985). Activated satellite cells were not detected in damaged muscle before phagocytosis of the necrotic debris by macrophages had begun. This finding clearly demonstrates that the removal of the necrotic muscle debris by macrophages seems to be a prerequisite for the activation and proliferation of satellite cells.

Diverse mechanisms have been assumed to be involved in the activation of satellite cells, including events of the inflammatory response and the release of growth factors.

VCAM-1, called very late antigen 4 (VLA-4), is present in leucocytes infiltrating muscle early after injury. Cell–cell interactions of infiltrating VLA-4(+) leucocytes and satellite cells, mediated by VCAM-1/VLA-4 ligation, mediates the recruitment of leucocytes to muscle after injury and focuses the invading leucocytes specifically to the sites of regeneration (Jesse et al. 1998). Infiltrating leucocytes may also initiate a series of molecular events implicated in muscle regeneration. On the other hand, the activation of satellite cells seems to be dependent on factors either synthesized and secreted by macrophages, or released from the necrotic tissue by phagocytes (Hurme & Kalimo 1992). These soluble factors have not been clearly identified but could be growth factors released either by neutrophils or invading macrophages. Several cytokines and growth factors have been shown to play key roles in the activation and proliferation of satellite cells (Table 3.1) (Husmann et al. 1996).

Growth factor and cytokine responses during myofibre regeneration

Fibroblast growth factors

The actions of FGFs on muscle precursor cell activation have been extensively studied (Sheehan & Allen 1999). FGFs exist as nine different isoforms but only FGF-1, -2, -4, -6 and -9 are active on satellite cells. In normal skeletal muscle, FGFs are stored in the extracellular matrix and are bound to proteoheparan sulphates. After muscle injury, FGFs are released, and high levels of both mRNA and protein have been reported in proliferating satellite cells before the formation of myotubes. High concentrations of FGF-1 have been previously correlated with a high level of myoblast proliferation in damaged muscles of dystrophin-deficient mice that display persistent cycles of myofibre degeneration–regeneration (Dimario & Strohman 1988).

It is thus very likely that FGFs contribute to the proliferation of satellite cells and are involved in the chemotaxis of further muscle precursor cells. The main action of FGFs is the activation
of muscle precursor cell proliferation to provide enough cells to allow regeneration to take place. Studies using mouse knockout models show the ability of redundant FGF family members to compensate for one another to preserve the role of these growth factors on satellite cell proliferation under pathological conditions.

Moreover, the availability of basic FGF, also called FGF-2, is of importance for capillary growth during muscle regeneration—a key factor for the success of muscle repair. Another member of the FGF family, FGF-5, might be important for the reinnervation process, when the presence of nerve is important for recovering the expected contractile and metabolic phenotype. Even bFGF can interact with other growth factors to stimulate the secretion of nerve growth factor (NGF) and contribute to the survival of neurones (Unsicker et al. 1992).

Collectively, FGFs are mainly involved in the activation and proliferation of satellite cells, in order to provide and activate stem cells to allow regeneration to occur. The exact mechanisms by which FGFs push satellite cells towards proliferation remain unknown, but it could be hypothesized that FGFs promote satellite cell activation, at least partly, by inhibiting the differentiation of myoblasts into myotubes. This effect of FGFs could be mainly related through inhibition of the expression of insulin-like growth factor II (IGF-II) (Rosenthal et al. 1991) or MyoD, one of the mammalian MRFs (Vaidya et al. 1989). As a consequence, other factors must either repress FGF production and/or strongly promote the differentiation of myoblasts, later during recovery, in order to minimize the repressing activity of FGFs on the differentiation of muscle precursor cells into myoblasts.

During the period of satellite cell proliferation, the expression of FGF receptor 1 (FGF-R1) increases, and the decrease in FGF-R1 expression is associated with a concomitant increase in satellite cell differentiation.

Platelet-derived growth factor
PDGF is first released from degranulated platelets, and later by activated macrophages and vessels in damaged muscle. This growth factor is chemotactic for adult muscle precursor cells and shows a highly stimulating effect on the proliferation of satellite cells (Ross et al. 1986). Moreover, together with FGFs, PDGF exerts a strong inhibition of the terminal differentiation of satellite cells into myoblasts.

Transforming growth factor beta
Like PDGF, TGF-β is released from degranulating platelets after injury. This growth factor is chemotactic for macrophages and leucocytes, contributes to the synthesis of proteins of the extracellular matrix, and plays a pivotal role in angiogenesis. Many of the biological effects of TGF-β concern the reorganization of the extracellular matrix, particularly the reconstruction of the basement membrane surrounding the regenerating myofibres. Those effects are mediated by the production of: (i) extracellular matrix proteases; (ii) protease inhibitors such as plasminogen activator inhibitor 1 (Laiho et al. 1986); and (iii) extracellular matrix components. TGF-β plays an important role in regulating repair after muscle injury and it has been suggested that excessive TGF-β-induced deposition of the extracellular matrix can lead to fibrosis (Border & Ruoslahti 1992). TGF-β regulates a subset of genes that encode growth factors and their receptors, and this finding could help to explain the varied cellular responses to TGF-β (Nielsen-Hamilton 1990). As for FGFs, TGF-β can be stored in an inactive form in the extracellular matrix and thus both are available for direct action after injury, without the need for new synthesis before their local action. Moreover, this growth factor exerts a control on satellite cells by inhibiting their proliferation and terminal differentiation, mainly by silencing the transcriptional activation of the MyoD family members (Allen & Boxhorn 1989).

Other cytokines
Other cytokines such as IL-6 and leukaemia inhibitory factor (LIF) have been shown to stimulate the proliferation of myogenic precursor cells in
vitro (Kurek et al. 1996). LIF is markedly increased very early after muscle injury, probably prior to infiltration of immune cells. The activating effect of LIF on the proliferation of myogenic precursor cells is additive to those of other growth factors such as bFGF and TGF-β, which suggests different mechanisms of action (Husmann et al. 1996). LIF has been shown to be produced by both damaged muscle fibres and resident non-muscle cells such as leucocytes and macrophages (Kurek et al. 1996).

IL-6 appears later in damaged muscle, between 12 and 24 h after injury. This cytokine, together with LIF, is one of the putative growth factors secreted by macrophages, which accounts for the role of these infiltrating monocytes on the proliferation of myogenic precursor cells (Cantini et al. 1994). As discussed above, there is clear evidence that besides their scavenger role, macrophages play a pivotal role in the activation and proliferation of satellite cells and myogenic precursor cells during muscle regeneration. Moreover, IL-6 promotes the degradation of necrotic tissue and induces apoptosis of macrophages following muscle injury (Cantini & Carraro 1996).

**Insulin-like growth factors**

Another group of growth factors, IGF-I and IGF-II, also called somatomedins, has been shown to be mitogenic for satellite cells and myoblasts (Florini et al. 1996). Besides these effects on the proliferation of myogenic cells, IGF-II also shows a marked stimulating effect on myoblast differentiation.

**Other growth factors**

Many other growth factors are already known to play a role in the regulation of skeletal muscle regeneration. Heparine-affine regulatory peptide, ciliary neurotrophic factor, NGF and hepatocyte growth factor (HGF) can interact with the most potent growth factors discussed above, to contribute in the activation of satellite cells, their proliferation and the reconstruction of the basement membrane and extracellular matrix (Husmann et al. 1996). HGF and its receptor c-Met have been localized in satellite cells and adjacent myofibres, and their expression has been related to the extent of muscle injury (Hawke & Garry 2001). Furthermore, HGF has been involved in the activation of satellite cells, as well as in the inhibition of myoblast differentiation, mainly by an inhibition of the expression of myogenic regulatory factors. Clearly, the efficiency of muscle regeneration is related to the fine growth factor interactions and to their role in the expression of myogenic regulators and extracellular matrix components.

Furthermore, differentiation of myogenic precursor cells has been shown to be regulated by a family of regulatory factors that can interact with certain of the growth factors mentioned above. The molecular mechanisms involved in the control of the differentiation of myogenic precursor cells and the role of MRFs should be examined.

**Differentiation of myogenic precursor cells**

**GROWTH FACTORS AND MUSCLE CELL DIFFERENTIATION**

Most growth factors involved in the activation and proliferation of satellite cells inhibit their differentiation into myoblasts, except for IGFs which exert a stimulating effect on the differentiation and fusion of myogenic cells (Table 3.1). It has been shown that IGF-I exerts less activation of myoblast differentiation than IGF-II, but taken together these growth factors play a key role at the onset of differentiation (Husmann et al. 1996). Some experimental evidence suggests that IGF-I acts first by stimulating muscle precursor cell proliferation, subsequently activating muscle cell differentiation. As specified above, several local growth factors inhibit muscle cell differentiation; this is the case for FGFs and it has been suggested that this inhibitory effect is related to a down-regulation of the IGF-II gene. It appears that a strong interaction exists between several growth factors such as FGFs, IGF-I and IGF-II.
MRFS AND MUSCLE CELL DIFFERENTIATION

Myogenic regulatory factors, members of the basic helix–loop–helix family (bHLH), are expressed in the embryo during development. These molecular factors activate the myogenic programme in embryonic precursor cells, thus initiating muscle differentiation. This family of myogenic factors, which includes MyoD, Myf5, MRF4 and myogenin, forms dimers with ubiquitous proteins. The heterodimeric complexes are transcription factors that bind to the E-box consensus DNA sequence that is found in the regulatory region of many muscle-specific genes. Studies support a role for MyoD and Myf5 in the determination of the myogenic cell fate and the formation of myoblasts, while myogenin and MRF4 appear to act later on muscle differentiation (Hawke & Garry 2001). Thus, these myogenic factors play a pivotal role in the commitment and differentiation of embryonic myoblasts during development (Seale & Rudnicki 2000).

The MRF expression programme during regeneration is similar to that observed during embryonic development. Not detected in quiescent satellite cells, MyoD is expressed very early, within 12 h after muscle injury, prior to molecular markers of cell proliferation (Mendler et al. 1998; Seale & Rudnicki 2000). The expression of MyoD mRNA has been strongly related to the proliferation of satellite cells and the beginning of regeneration (Olson & Klein 1994). Satellite cells entering the cell cycle express first either mainly MyoD or more seldom Myf5, followed by the coexpression of these two factors. The expression of these two early acting factors is shown to have a similar time course in regenerating slow and fast muscle. Following proliferation, myogenin and MRF4 are expressed in cells during their differentiation and expression of muscle-specific proteins. However, the expression of these late-acting MRFs changes in a different way in regenerating fast and slow muscles (Mendler et al. 1998).

Levels of myogenin and MRF4 mRNA decrease transiently in slow muscle, while mRNA levels remain relatively constant in fast-twitch muscles. MRF mRNA is lacking in quiescent satellite cells and this finding is consistent with the hypothesis that these cells represent stem cells, distinct from myoblasts, that give rise to myogenic cells after the beginning of their proliferative phase. In summary, an up-regulation of MyoD and Myf5 appears to be required for satellite cells to enter the proliferative phase, while MRF4 and myogenin appear later to participate to the regulation of their differentiation into myoblasts.

Skeletual muscle maturation

The differentiation of myogenic precursor cells is characterized by the capability for myoblasts to synthesize muscle-specific proteins. Differentiated myoblasts then migrate side by side, loose their membranes and fuse to form multinucleated myotubes. The fusion of myoblasts to form myotubes is associated with the expression of muscle-specific proteins. Striking changes of gene expression occur with the fusion of myoblasts, and genes encoding muscle-specific proteins are turned on.

The presence of specific proteins in the tissue is required to produce a viable cellular structure and/or to play a tissue-specific role accounting for the properties of a system. Thus, skeletal muscle can be approached from a molecular viewpoint, particularly the contractile proteins which account for the mechanical function and performance. Several families of proteins present a high degree of molecular variability, due to the existence of multiple isoforms. Some of the proteins accounting for the contractile and metabolic properties exist as adult mature and developmental immature isoforms.

Skeletal muscle is characterized by the existence of numerous types of fibres and a highly organized arrangement, resulting in a variety of functional capabilities. A high molecular variety of myofibrillar and enzymatic proteins, due to the existence of multiple isoforms, accounts for this myofibre diversity. Three isozymic systems are particularly relevant to the functional heterogeneity of myofibre types: myosin and especially the isoforms of myosin heavy chain (MHC), the
creatine kinase (CK) isozymes and the lactate dehydrogenase (LDH) isozymes.

**Myosin transition in regenerating skeletal muscle**

One of the most informative methods of delineating muscle fibre types is based on the examination of specific myosin profiles. Myosin is the most essential part of the contractile machinery and contributes to the functional diversity of myofibres. A strong correlation has been shown between myosin isoforms, contractile properties and physiological measurements of myofibres (Schiaffino & Reggiani 1996). The myosin isoform composition is thus a molecular marker for different muscle fibres, for the physiological state of muscle tissue, and for the maturation state of regenerated muscle.

It is a well-known fact, that muscle regeneration recapitulates the embryonic development of skeletal muscle fibres. The new myogenesis that takes place in regenerating muscles has been related to the myogenesis which occurs during normal development. The major steps of regeneration closely recapitulate those of normal ontogenesis, and these two types of myogenesis display many similar morphological and biochemical features (Swynghedauw 1986). However, the transition from the immature to adult isoforms is more precocious in regenerating than in developing muscles (d’Albis et al. 1988).

The sarcomeric myosin molecule is a hexamer, consisting of two MHCs, two essential or alkali myosin light chains (MLCs) and two regulatory or phosphorylatable MLCs. MHCs account for about 85% of the myosin molecule. This subunit has been found to be both highly conserved in structural terms and highly polymorphic (i.e. present as numerous isoforms). Two developmental isoforms have been described in rat and human skeletal muscle: MHC-emb, predominant in embryonic stages, and MHC-neo, predominant in perinatal stages (Whalen et al. 1981). Rat skeletal muscles contain three major fast MHC isoforms, called MHC-IIa, MHC-IIx and MHC-IIb, and one slow isoform, called MHC-β/slow, that is present in slow fibres (Schiaffino & Reggiani 1996).

**MHC expression during development**

Muscle development is characterized by the asynchronous differentiation of successive fibre generations—into at least primary and secondary generations of myofibres in rats. These sequential events have been mainly described in rats and mice, since in these species a more complete picture of MHC isoform expression is available. At early stages of embryonic development, all fibres express high levels of MHC-emb and low levels of MHC-neo and MHC-β/slow. By days 16–17, MHC-β/slow and MHC-neo are segregated in different myofibres, and fibre-type diversification is then detectable (Condon et al. 1990). At this stage of muscle development, primary generation fibres express either MHC-emb and MHC-β/slow, or MHC-neo and MHC-emb. Secondary generation fibres express mainly MHC-neo.

The major phase of myofibre maturation occurs during the perinatal period, with the appearance of the adult fast MHC isoforms and the emergence of the definitive fibre types. Myofibres expressing MHC-emb and MHC-neo undergo a further diversification process during the postnatal week in rats. They give rise to one of the four major fibre-type phenotypes, characterized by the expression of at least one of the four adult MHC isoforms (Schiaffino & Reggiani 1996). These adult MHC isoforms are initially coexpressed with MHC-emb and/or MHC-neo, which progressively disappear during the first month of life. The MHC isoform profile of myofibres is not definitively established during the early postnatal period but undergoes further changes with ageing.

**MHC isoform transitions in regenerating muscles**

New myofibres formed after muscle injury transiently express developmental myosin isoforms before switching to adult isoforms. Immature
Different types of myosin are synthesized before a normal adult pattern is achieved (Sartore et al. 1982; Whalen et al. 1990). However, embryonic and neonatal isoforms disappear faster during regeneration than during normal myogenesis (d’Albis et al. 1988). Moreover, both fast- and slow-twitch regenerating muscles show the same transition, first toward a predominantly fast-type isoform profile, and secondly toward a slow-type profile for slow-twitch muscles.

It is not clear to date if satellite cells are programmed to express specific myosin isoforms. Some data suggest that satellite cells are not predetermined with respect to the type of adult myosins which will accumulate in the fibres they form (Whalen et al. 1990). This finding supports the notion that the programme expressed by satellite cells is not strictly determined and could be influenced by several factors, including the presence of nerve. More recent data clearly point to intrinsic differences between satellite cells from fast and slow muscles (Düsterhoft & Pette 1993; Dolenc et al. 1994; Martelly et al. 2000).

Two major controlling influences play a role in determining the specific adult myosin type that ultimately appears in mammalian muscle fibres. Continuous innervation is an important factor for the maintenance of slow myosin expression during regeneration in rats (d’Albis et al. 1988; Whalen et al. 1990). Innervation plays a clear role in determining which adult myosin isoform will accumulate in regenerated muscle. Slow myosin is expressed in regenerated slow-twitch muscles only in the presence of slow nerves. In contrast, a switching of myosin from embryonic and neonatal MHC isoforms toward fast, and not slow, isoforms was observed in regenerated slow muscle in the absence of nerve, suggesting the existence of a ‘default programme’ of MHC expression (Buttler-Browne et al. 1982).

Thyroid hormone levels also contribute to the control of MHC isoform transitions during regeneration. Hypothyroidism has been shown to inhibit the replacement of neonatal myosin by adult fast myosin isoforms during regeneration; conversely, hyperthyroidism induces a precocious induction of fast myosin expression (d’Albis et al. 1987). It is thus clear that the presence of motor nerves and the thyroid hormone status play determinant roles in myosin isoform expression, with potential consequences on muscle function.

Members of the bHLH family of transcriptional regulators probably play a role in the expression of the MHC isoforms during regeneration. The E-box sequence, specific to the heterodimers made up of MRFs and E-proteins (see above), has been found in the regulatory regions of several genes encoding for phenotypic proteins, including fast MLCs and type I and type IIB MHCs (Talmadge 2000). The primary function of MRFs is to initiate the expression of muscle-specific proteins during development and regeneration. Moreover, it has been suggested that myogenin and MyoD could play a role in determining the slow and fast phenotype, based on the distribution of these two factors in slow and fast myofibres, respectively (Hughes et al. 1993).

### Creatine kinase in regenerating skeletal muscle

The enzyme CK has been involved in the maintenance of the intracellular energy supply of cells with intermittently high and fluctuating energy requirements, such as skeletal muscle. It has been suggested that CK, which catalyses the reversible reaction:

\[
\text{Phosphocreatine}^{2-} + \text{MgADP}^- + H^+ \leftrightarrow \text{Creatine} + \text{MgATP}^{2-},
\]

fulfils different roles in fast- and slow-twitch muscles. This enzyme exists as multiple isoforms; cytosolic isoforms of CK are heterodimeric associations of two types of monomer subunits, known as M, muscle, and B, brain (MM-, MB- and BB-CK). The MM-CK isozyme is bound to myofilaments and rephosphorylates ADP produced by myosin ATPase, contributing to maintain a locally high ATP:ADP ratio during periods of muscle contractions. Moreover, two additional CK isozymes have been detected in mitochondria, either specific for striated muscle, called \(m_i^a\)-CK, or ubiquitous, \(m_i^b\)-CK. These isozymes
play pivotal roles in high oxidative muscles by favouring phosphocreatine resynthesis from oxidative phosphorylation (Ventura-Clapier et al. 1994). The activity and pattern of this enzyme is of functional relevance since resistance to fatigue is partly related to the ability of skeletal muscle to sustain ATP levels through increased oxidative metabolism.

During myogenesis, there is a progressive increase in the total activity of CK, as well as a shift from BB- to MM-CK isozyme. The onset of the expression of the M-subunit of CK and the increase in the specific activity of the MM-CK isozyme have been correlated with the development of functional muscle. Moreover, the M-CK promoter is under the control of MyoD, one of the most early MRFs, and the expression of the M-subunit is associated with that of other muscle-specific protein isoforms encoded by genes having an E-box sequence in their regulatory regions. The developmental pattern of expression of the various isozymes has been studied in the heart, where this multienzyme system appears to be essential for the adaptation to increased metabolic demand (Hoerter et al. 1994). However, little is known about the expression of mi-CK during the development of slow oxidative skeletal muscle.

Skeletal muscle degeneration using bupivacain is associated with a marked decrease in the total activity of CK (Plaghki 1985), and an increase in the specific activity of the BB-CK isozyme. The CK isozyme composition recovers a pattern similar to that of control muscle 10 days after initial injury (Sadeh et al. 1984). The specific activity of mi-CK isozyme, linked to the site of production of energy, is similar in control and regenerated skeletal muscle 8 weeks after infiltration with a myotoxin (Bigard et al. 2000). However, the time-course of expression of mi-CK during regeneration remains to be examined. This issue is of importance in slow oxidative muscles because mi-CK located in the inner mitochondrial membranes couples ATP synthesized during oxidative phosphorylation with creatine formed during muscle metabolism to produce phosphocreatine.

Lactate dehydrogenase activity in regenerating skeletal muscle

Lactate dehydrogenase exists in multiple molecular forms as a tetrameric association of two types of monomer subunits, known as H (heart) and M (muscle), which may be combined in five different ways and result in five isozymes. A correlation between the LDH isozyme pattern and the anaerobic/aerobic glycolysis capacity of muscle has been suggested. Heart contains a predominantly H type of LDH subunit, while fast glycolytic muscle contains the M type. The LDH isozyme pattern of slow oxidative skeletal muscles is roughly mixed with equal specific activities of H- and M-subunits.

During the fetal development, there a shift from the H toward the M-subunit, while the specific activity of the M-subunit increases during postnatal development (Plaghki 1985). A marked decrease in the total activity of LDH has been reported early after muscle injury. Thereafter, the activity of total LDH increases from the fusion of precursor myogenic cells to myotubes. As for the expression of MHC isoforms, regeneration recapitulates the events that normally occur during myogenesis, and the pattern of LDH isoforms shifts from a predominance of H-subunits toward M-subunits in regenerated slow muscles. Regenerated fast skeletal muscles recovered a LDH isoform pattern similar to that expected by 2 months after initial injury (Bigard et al. 2000).

Practical consequences for muscle healing

Treatment of muscle injuries varies widely depending on the nature and severity of the trauma. Besides the classic treatment of injured muscle, different biological approaches can be used to enhance muscle healing. Recent advances in knowledge of the mechanisms of muscle degeneration and regeneration after injury highlight the functional importance of the inflammatory response and support the therapeutic potential of several growth factors to improve muscle
healing. On the other hand, the physical activity pattern is an environmental factor known to affect muscle growth and maturation during the early phases of regeneration. Therefore, it is of clinical interest to determine the specific effects of exercise on muscle healing.

**Initial inflammatory response to muscle injury**

As discussed above, the injured muscle first undergoes necrosis and is infiltrated by macrophages and neutrophils. Inflammatory cells are important components of the successful repair of injured muscle. The early use of NSAIDs, which inhibit the function of macrophages and neutrophils, leads to a slower muscle regeneration than in untreated animals (Almekinders & Gilbert 1986). The impaired recovery in the muscles of treated animals was attributed to both the slow removal of cellular debris and the decrease in synthesis and release of soluble factors important for the activation of regenerative processes by macrophages. Because the early response of inflammatory cells to muscle injury is a determining factor in directing muscle repair, there is an argument for not using anti-inflammatory drugs during this period. It appears especially important to respect the inflammatory response to acute muscle injury, with the aim of protecting the macrophage invasion, first to remove cellular debris and second to promote the involvement of activated macrophages in muscle fibre regeneration by producing specific soluble factors. Whether the rest, ice, compression and/or elevation components of the principal therapy affect the inflammatory response to acute muscle injury has not been determined to date. Immobilization after muscle injury has been shown to limit the amount of connective tissue, but it is likely that this beneficial effect is not directly related to the inflammatory response.

**Growth factors and muscle healing**

The activation of satellite cells, and their proliferation, differentiation into myoblasts and fusion to form myotubes and myofibres are under the control of many soluble factors such as growth factors (see above). Many growth factors interact to control the involvement of inflammatory and satellite cells in the healing process, and to regulate the reorganization of the basement membrane and the extracellular matrix. Several studies have previously shown that individual or combined growth factors play specific roles during regeneration and therefore are able to improve muscle healing. The number and the mean diameter of regenerating fibres increased in muscles receiving serial and direct injections of bFGF and IGF-I in comparison with non-treated muscles, reflecting improved healing (Fig. 3.5) (Kasemkijwattana et al. 2000; Menetrey et al. 2000). At 1 month, and in contrast with non-treated muscles, regenerating myofibres were uniformly

![Fig. 3.5 Changes in the mean diameter of regenerating myofibres when basic fibroblast growth factor (bFGF) and insulin-like growth factor I (IGF-I) are used to improve muscle healing after strain injury ($P < 0.05$). (From Kasemkijwattana et al. 2000.)](image-url)
located in the deep and superficial parts of the muscle treated with bFGF and IGF-I, their mean diameter was similar to the surrounding normal fibres, and many of their nuclei were already peripherally located. Moreover, the development of fibroblastic tissue was also reduced in treated muscles, suggesting that muscle healing was accelerated in these muscles when compared with non-treated muscles. A significant improvement in fast-twitch and tetanic strength was observed 15 days after injury in muscles treated with growth factors (bFGF and IGF-I) when compared with sham injected injured muscles (Fig. 3.6). Taken together, these studies demonstrated that selected growth factors properly applied and injected within injured muscle are able to enhance muscle healing after injury.

However, it is unlikely that a treating physician or a patient will accept repeated intramuscular injections or implanted perfusion pumps, particularly with the infection risk. Thus, other approaches have been considered to achieve a sustained expression of exogenous genes to skeletal muscle, such as myoblast transplantation or gene therapy based on viral and non-viral vectors. The application of these emerging technologies to deliver sustained expression of growth factors in the injured muscle has been previously examined (Kasemkijwattana et al. 1998). In this study, the ability of recombinant adeno-associated virus to mediate direct and \textit{ex vivo} the transfer of a marker gene (β-galactosidase) within the contused injured muscle, suggested that this biological approach could be able to deliver an efficient and persistent

![Bar charts](image-url)
expression of selected growth factors. It is likely that the development of those vectors carrying selected growth factors and the delivery of growth factors, using such new technologies, represent potential strategies in improving muscle healing that will be available for humans in the near future.

**Heparin sulphate proteoglycans and muscle repair**

Growth factors such as FGFs and TGF-β are transmitted to the cell either by one tyrosine kinase receptor or by binding to low-affinity heparan sulphate proteoglycans located on the cell surface and in the extracellular matrix. As for other growth factors, the interaction between them and the heparan sulphate components of the extracellular matrix might play a pivotal role in regeneration. Selected dextran derivatives have been shown to mimic some properties of heparin and heparan sulphate to stabilize and protect heparin-binding growth factors such as FGFs and TGF-β. A single injection of a dextran derivative was found to improve the regenerative process in a fast skeletal muscle after a crush injury (Gautron et al. 1995). Besides clear positive effects on the histological structure of regenerated muscle, dextran derivatives were shown to accelerate the shift from the neonatal to adult MHC isoforms (Aamiri et al. 1995). Crushed slow muscle that had regenerated in the presence of dextran derivatives showed a fourfold increase in the number of slow fibres, in comparison with muscle regenerating in the absence of dextran derivatives (Fig. 3.7). It is thus likely that dextran derivatives act by protecting and favouring the effects of heparin-binding growth factors (FGFs and TGF-β) involved in the process of muscle regeneration. This family of polymers may therefore open a new therapeutic approach to accelerate skeletal muscle repair after injury.

**Physical exercise and muscle regeneration**

To determine the impact of physical exercise and contractile activity on muscle regeneration requires the use of animal models. A decreased mechanical load after the removal of weight-bearing activity impairs the rate and degree of growth and maturation during regeneration (Esser & White 1995). The degree of restoration to control adult protein concentration values of regenerating muscle is markedly altered by decreased mechanical load. Even the timing of transitional expression of MHC isoforms is delayed in unweighted regenerating muscles (Esser & White 1995; Bigard et al. 1997). There are thus experimental data suggesting that growth and maturation are impaired by a decreased mechanical load on regenerating muscles. On the other hand, increased mechanical load by
ablation of synergistic muscles enhances the growth of regenerated muscle but not maturation (Esser & White 1995). While ablation of synergistic muscles is a model of chronic mechanical overload, it is of interest to evaluate the specific influence of the metabolic component by examining the response of regenerating muscle to running exercise. The growth of damaged muscle and its oxidative capacity are improved by run conditioning. On the other hand, the response of the MHC composition to endurance training has been shown to be more marked in regenerated muscle, at least for fast MHC isoforms (Bigard et al. 1996). Taken together, these results clearly demonstrate that the removal of contractile activity and mechanical load have deleterious effects on muscle growth and maturation. It is thus suggested that muscular activity plays a key role in the recovery of damaged skeletal muscle.

Conclusions

Although studies using laboratory animals amply demonstrate that skeletal muscle can regenerate after injury, our knowledge of muscle regeneration in humans is fragmentary and is still growing to date. The treatment and prognostics of muscle injury vary widely according to the severity and the extent of the trauma, more particularly according to the presence of nerve and vascularization. Satellite cells, normally in a quiescent state, are stem cells activated in response to muscle injury, giving rise to myogenic cells, myotubes and new myofibres. In addition, non-muscle cells also play complex and essential roles in regulating the muscle repair response, in removing cellular debris, in activating the proliferation of myogenic cells, and in initiating structural changes in myofibres and extracellular matrix. The time-course changes in the expression of muscle-specific proteins, such as MHC, CK and LDH, should be considered during muscle regeneration, to control the recovery of the functional properties of skeletal muscle.

Many different growth factors interact simultaneously and sequentially with satellite cells to trigger muscle-specific gene products. Most of these growth factors stimulate the proliferation of satellite cells and concomitantly inhibit their differentiation and fusion, except for IGFs, which promote the differentiation of the myogenic precursor cells. TGF-β also plays an important role in muscle regeneration in promoting the reorganization of the basement membrane and extracellular matrix components. The interaction of those growth factors during muscle regeneration requires further investigation. A better understanding of growth factor interactions and their role in the expression of myogenic regulators and extracellular matrix components is needed to identify specific growth factors or cytokines able to improve skeletal muscle repair. Exercise resting with increased contractile activity but without mechanical overload is good advice in promoting efficient muscle healing after damage. On the other hand, the improvement of muscle healing after injury may be enhanced by the development of new technologies, such as gene therapy, that may be able to deliver target proteins to the damaged muscle without systemic side effects.

References


Skeletal Muscle Regeneration


Introduction

This chapter will briefly review recent basic scientific information on the structure and biomechanics of tendons and ligaments, discuss mechanisms of injury to ligaments/tendons and their repair response, and attempt to relate this to practical clinical patient rehabilitation management. Several recent, excellent reviews are available in this area: Benjamin et al. (1986, 1995), Daniel et al. (1990), Kannus et al. (1992a, 1992b), Jackson et al. (1993), Archambault et al. (1995), Benjamin and Ralphs (1995), Viidik (1996), Frank and Jackson (1997), Jozsa and Kannus (1997), Frank (1999), Frank et al. (1999), Rodeo and Izawa (1999) and Woo et al. (2000).

Structure and biomechanics of ligaments and tendons

There are subtle differences between ligament and tendon morphology (Amiel et al. 1983) but for the sake of this discussion they will be treated as very similar tissues.

The fibroblasts of mature tendons lie in longitudinal rows and are flat, very elongated cells squeezed laterally between the collagen fibrils. Recently McNeilly and associates (1996) have demonstrated elegantly (using confocal microscopy coupled with fluorescent dye cellular labelling techniques) that tendon fibroblasts communicate with one another via an extensive three-dimensional network of long cell processes and gap junctions ramifying between the collagen matrix (Fig. 4.1). This new information indicates the possibility of intercellular ‘talk’ between cells similar to that of osteocytes and hence that these tendon cells may be able to sense and coordinate a response to mechanical load or lack of load.

Mature adult ligament and tendon are composed of large diameter type I collagen fibrils (150 nm or more in diameter) tightly packed together in a rope-like configuration with a small amount of type III collagen dispersed in an aqueous gel containing small amounts of
The repair of tendons and ligaments involves proteoglycan and elastic fibres. (Figs 4.2 and 4.3). The outstanding feature of both these unique load-bearing tissues is the collagen ‘crimp’, which is a planar wave pattern found extending in phase across the width of all ligaments and tendons (Fig. 4.4) (Diamant et al. 1972). This collagen ‘crimp’ appears to be built into the tertiary structure of the collagen molecule and is probably maintained in vivo by inter- and intramolecular collagen cross-links as well as a strategically placed elastic fibre network. The ‘crimp’ may help to attenuate muscle loading forces at the tendoperiosteal junction as well as at the musculotendinous junction.

Viidik (1996) has recently reviewed the basic biomechanics of ligaments and tendons and the reader is referred to this text for a more detailed examination of structure–function relationships. The basic function of tendons and ligaments is to transmit force with ‘reasonable safety margins’ (Viidik 1996). The longitudinal array of collagen fibrils converts externally applied axial loads to internal lateral compression which results in interfibril friction and heat generation. This heat generation may be responsible for central fibroblast death in thick tendons, which has recently been described as exercised-induced hyperthermia (Wilson & Goodship 1994). Peak intratendinous core temperatures in the range of 43–45°C have been recorded in race horse tendons. Temperatures greater than 42.5°C are known to cause fibroblast death in vitro.

Also of interest is the notion that the muscle–tendon unit can act as a spring and the tendon can recoil after being eccentrically stretched, for example during the landing phase of a hopping kangaroo. During this landing phase the tendon stretch of the Achilles tendon is stored as elastic energy which is then converted to add extra lift for the animal during the concentric phase of the hop and hence leads to energy conservation (Morgan et al. 1978; Proske & Morgan 1987). Indeed, in the slowly hopping wallaby, strain energy storage in tendons and ligaments accounts for 33% of the negative and positive work that is done whilst the feet are on the ground (Ker et al. 1986)!

---

**Fig. 4.2** Adult rat ACL collagen fixed with ruthenium red after treatment with elastase at pH 8.8. for 12 h. There is typical regular banding periodicity (labelled) and linking filaments probably hyaluronate, which link the fibrils between the c and d bands. The fibre marked with the asterisk appears to be dividing into a smaller fibril. Proteoglycan granules are attached to the fibrils in the region of the c and d bands. (Magnification × 98 000.)

**Fig. 4.3** Approximate amounts of main tendon/ligament components by weight. Collagen forms the main protein component of the dry weight with smaller amounts of elastin and proteoglycans.
Ligament and tendon injury can be closely correlated with the load–deformation curve (Butler et al. 1979; Oakes 1981). The load–strain curve can be divided into three regions (Fig. 4.5):

1. The ‘toe’ region or initial concave region represents the normal physiological range of ligament/tendon strain up to about 3–4% of initial length, and is due to the flattening of the collagen ‘crimp’. Repeated cycling within this ‘toe region’ or ‘physiological strain range’ of 3–4% (which may be up to near 10% in cruciate ligaments due to the intrinsic macrospiral of collagen cruciate fibre bundles) can normally occur without irreversible macroscopic or molecular damage to the tissue.

2. The second part of the load–deformation curve is the linear region where pathological irreversible ligament/tendon elongation occurs due to partial rupture of intermolecular cross-links. As the load is increased further intra- and intermolecular cross-links are disrupted until macroscopic failure is evident clinically. Electron microscopic studies (Viidik & Ekholm 1968) have shown the collagen fibrils are elongated in this phase and the periodicity increases from 67–68 nm to 72 nm and the banding pattern may become disjointed across the fibrils indicating intrafibril damage by molecular slippage under shear strain (Kastelic & Baer 1980). With continued loading into the third region, the banding pattern is completely lost and in immature collagen no free ends are seen ultrastructurally. This suggests that damaged fibrils can still bear
load and may represent a form of work hardening. In mature tendon, ruptured collagen fibril ends can be observed. Elegant X-ray diffraction and electron microscopic studies, coupled with accurate fast and slow loading studies, of both rat tail and human finger tendons has confirmed that initial damage to the fibril is an intrafibrillar sliding process that occurs only a few milliseconds before macroscopic fibre slippage occurs (Knorzer et al. 1986). The early part of the linear region corresponds to mild ligament tears or grade 1 ligament tears (0–50% fibre disruption) and the latter part to grade 2 (50–80% fibre disruption) where there is obvious clinical laxity on stress testing. Grade 1 and 2 injuries always have some pain after the initial trauma, and usually with a grade 2 injury the athlete cannot continue and this is a rough guide to the clinical severity of the injury.

3 In the third region, if continued loading occurs, the linear part of the curve flattens and then the yield or failure point is reached at 10–20% strain, depending on the ligament/tendon fibre bundle macro-organization. In this region, complete ligament tendon rupture occurs at ‘maximal breaking load’ and this is the dangerous grade 3 ligament rupture seen on clinical testing. It is ‘dangerous’ because the athlete has severe momentary pain when the trauma is applied and then little pain, and the athlete (and often inexperienced examiners) erroneously believe the injury is trivial and treat it as such with disastrous consequences.

Tendons and ligaments behave as non-linear elastic materials. Viscous and plastic behav-our become obvious when cyclical loading–unloading protocols are used during laboratory tensile testing. The laboratory testing protocols can be similar to the cyclical loading imposed on tendons and ligaments during running. During such cyclical testing the load–strain curves initially have a large hysteresis loop and as the cycles progress with time this hysteresis loop is lost; this is thought to be due to squeezing of water from the tissue and concurrently aligning the collagen fibrils along the axis of loading. This ‘preconditioning’ cycling could be thought of as an important part of the warm-up procedure in athletes and is reversible during the non-running period as water and proteoglycans redistribute amongst the non-loaded fibrils. If, whilst cycling testing a specimen of tendon in the laboratory, the cycling is stopped two further protocols can be used. If the stress is kept constant by slowly increasing the strain, a creep phenomenon results. If the strain or deformation is kept constant by increasing the load slowly, stress–relaxation of the tissue occurs (Viidik 1996).

An enormous amount of research has been performed on the biomechanical properties of the human knee ligaments, with the anterior cruciate ligament (ACL) dominating research because of its key role in anteroposterior stability of the knee. Rotational injury to the knee with the
foot fixed, as well as hyperextension of the knee, appear to be the two key mechanisms involved in human ACL disruption. Forces of the order of 2000 N are required to disrupt the ACL and are even higher for the posterior cruciate ligament (PCL). Direct falls onto the tibia or collisions with opponents, such that a posterior displacement force of the tibia occurs on the femur, appears to be a common mechanism for PCL injury. Colateral ligament injury involves excessive varus, valgus or rotational forces. The ligament–bone junction with its special fibrocartilage transition zone is a common region of clinical failure. Recent experimental animal studies indicate that ligament mid-substance strain is much lower than at the insertion sites and this appears to be due to a differing collagen-fibre crimp amplitude and angle. This higher strain at the insertion sites, together with ligament insertion geometry, could be an explanation of preferential failure of some ligament insertion sites—especially the femoral attachment of the medial collateral ligament of the knee, which has an almost $90^\circ$ insertion into the region of the medial femoral epicondyle compared with its tibial periosteal insertion.

**Ligament/tendon insertion to bone**

Benjamin and his co-workers (Benjamin et al. 1986, 1995; Benjamin & Evans 1990; Evans et al. 1990; Benjamin & Ralphs 1995) have recently studied extensively tendon–bone junctions where the fibrocartilage interface or enthesis exists between the tendon and the bone. The width of this unique fibrocartilage interface, they hypothesize, is dependent upon the relative degree of movement that occurs between the tendon and its bone attachment (Fig. 4.6). They suggest that the enthesis prevents collagen fibre bending and perhaps undergoing shearing, fraying and failure at this special junction region. They have found the thickest fibrocartilage interface in the very mobile Achilles tendon and a lesser thickness of this interface in tendons such as the tibialis anterior posterior and very little in the long flexor tendons (Frowen & Benjamin 1995). It is of interest that they observed longitudinal splits in the Achilles tendon with some evidence of repair and also transverse tears at the bone–tendon junction that were filled with fat cells and no repair tissue (Rufai et al. 1995), which Frank et al. (1999) recognize as mechanical ‘flaws’ during tendon repair (see below).

**Correlation of collagen fibril size with mechanical properties of tissues**

Parry et al. (1978) have completed detailed quantitative morphometric ultrastructural analyses of collagen fibrils from a large number of collagen-containing tissues in various species. They came to a number of conclusions that can be summarized as follows:
Type 1-orientated tissues, such as ligament and tendon, have a bimodal distribution of collagen fibril diameters at maturity.

The ultimate tensile strength and mechanical properties of connective tissues are positively correlated with the mass average diameter of collagen fibrils. In the context of the response of ligaments to exercise they also concluded that the collagen fibril diameter distribution is closely correlated with the magnitude and duration of loading of tissues (Fig. 4.7).

Collagen fibril diameter quantification with age and correlation with ACL tensile strength

Oakes (1988) measured collagen fibrils at various ages of the rat from 14 days fetal to 2-year-old senile adult rats. The mean diameter and the range of fibres from the largest to the smallest for each time interval were plotted against age and are shown in Fig. 4.7. The mean fibril diameter begins to plateau at about 7 weeks after birth. Also plotted on this figure is the separation force required to rupture the ACL in the rat with age. Special grips were used in this study to obviate epiphyseal separation and 70% of the failures occurred within the ACL. It can be seen that the two curves closely coincide, indicating a close correlation between the size of the collagen fibrils and the ultimate tensile strength of the ACL as has been suggested by Parry et al. (1978). This rapid increase in collagen fibril size over 6 weeks in the growing rat is not seen during normal ligament tissue repair or remodelling (Matthew et al. 1987; Oakes et al. 1991).

Work by Shadwick (1986, 1990) has also demonstrated a clear correlation between collagen fibril diameter and the tensile strength of tendons. He determined that the tensile strength of pig flexor tendons was greater than that of extensor tendons and that this greater flexor tendon tensile strength was correlated with a population of larger diameter collagen fibrils not present in the weaker extensor tendons. Also, studies by Oakes et al. (1998), using the rabbit patellar tendon, have demonstrated a high correlation between the area-weighted mean collagen fibril diameter (this method adjusts for the varying numbers of large and small fibrils within a tendon) and both modulus ($R = 0.79$) and ultimate tensile strength ($R = 0.63$) (Fig. 4.8).

It should be noted here that the original large collagen fibrils seen in normal tendons and ligaments are not replaced after ligament/tendon repair.
rupture or injury and this is the probable explanation for the poor material properties of repair scars even 1 year after repair. They are also not replaced when ligament replacements are used as scaffolds, such as for the reconstruction of the knee ACL when autograft or allograft tendons are used. This is a major explanation for the high failure rate of such knee ACL reconstructions, apart from any surgical misadventure (see below for further discussion).

Basic biomechanics of tissue injury

Muscle–tendon–bone injury

The basic causes of intrinsic muscle injury are still not entirely clear, but have been attributed to inadequate muscle length and strength (for example tight hamstrings, especially in adolescent boys), muscle fatigue and inadequate muscle skills. It is also clear that most muscle injuries occur in the lower limb and most involve two-joint muscles such as the hamstrings and the rectus femoris, probably because of the complex reflexes involved in simultaneous co-contraction and co-relaxation (Oakes 1984). An excellent review of current knowledge of muscle strain injuries has been published by Garrett (1996). Recently, Hartig and Henderson (1999) have clearly demonstrated that increased hamstring flexibility decreased lower limb overuse injuries in military basic trainees.

Both concentric and eccentric muscle–tendon unit loading can cause muscle–tendon–bone junction injury. The use of eccentric muscle loading to cause increased muscle hypertrophy, as against the use of more conventional concentric loading, has led to the phenomenon of eccentric

![Graph](image-url)
muscle soreness which is now known to be due in part to muscle sarcomere disruption at the Z-lines (Friden et al. 1983). Eccentric muscle–tendon–bone load can generate more force than concentric contractions and may be the mechanism by which the patellar tendon and its attachments lead to tendoperiosteal partial disruptions at both the superior and inferior poles of the patellar. Studies by Chun et al. (1989) demonstrated that the inferomedial collagen fibre bundles of the human patellar tendon, when subjected to mechanical analysis, fail at loads which are much less than the lateral fibre bundles. The biological reasons for this are not clear at the moment but it helps to explain the prevalence of inferomedial tenderness, which is such a common cause of anteromedial knee pain or ‘jumper’s knee’.

The bone–tendon junction or enthesis is one of the commonest sites for tissue injury, as is seen with the classic infrapatellar tendonitis which is so refractory to treatment. Benjamin et al. (1986) have suggested that the zone of fibrocartilage at the enthesis minimizes local stress concentrations and appears to be characteristic of tendons and ligaments where there is a great change in angle between the tendon or ligament and bone during movement. It is of interest that the enthesis which gives the most problems in the clinic (the inferior patellar pole and the patellar tendon) is the one with the least thickness of fibrocartilage (see Fig. 4.6). It is possible that a major mechanism of failure at this very mobile enthesis may be collagen fibril shear.

A study by Cooke et al. (1997) found that more than one-third of athletes with patellar tendinopathy were unable to return to sport for more than 6 months because of recurrent or persistent pain and eventually required surgery, indicating the difficulty of the reformation of a normal load-bearing enthesis after repeated tension–avulsion injury. Also it was clear from their studies that imaging such as ultrasonograms and magnetic resonance imaging (MRI) scans do not always correlate with the clinical and histopathological findings (Khan et al. 1998).

A recent detailed study of chronic patellar tendinopathy treated surgically by either open or arthroscopic patellar tenotomy, revealed that about 40% of subjects could not return to their previous level of sporting activity and that with either form of patellar tenotomy could expect little improvement in symptoms beyond 12 months postoperatively. In this study those patients that achieved postsurgical sporting success appeared to have an 80% chance of having prolonged success. There was no difference in outcomes between the open or arthroscopic procedures (Coleman et al. 2000).

**Muscle–tendon junction injury**

Failure at this junctional region is common. There is an increased muscle cell membrane folding of the terminal end of the last muscle sarcomere that has important mechanical implications for reducing the stress at this critical junctional region. It has been determined that a typical vertebrate fast-twitch cell can generate about 0.33 MPa of stress across the cell. The stress placed on the cell’s junctional complex at the muscle–tendon junction by the complex folding of the terminal sarcomeres experiences a maximal stress of $1.5 \times 10^4$ Pa, which is much less than $33 \times 10^4$ Pa, and this difference may determine whether mechanical failure occurs at this junctional region.

With muscle injury at this junctional site it is probable that the complex sarcomere muscle membrane infolding—to increase the surface area and hence decrease substantially the stress—is probably not reproduced following repair and this may be an explanation for the recurrence of tears at this junction in athletes. A detailed comprehensive review has been published by Noonan and Garrett (1992). Kannus et al. (1992a) (Kvist et al. 1991) have demonstrated in the rat that type II (fast-twitch) fibres have a more complex folding pattern than the type II (slow-twitch) fibres. The type II fibres have an increase of 30–40% in contact surface area compared to the type I fibres. In this way the larger forces transmitted via the myotendinous junction with type II muscle can be transmitted through this junction without increasing the force applied per surface unit of the junction (Fig. 4.9).
Tendon injury

Tendon injury in sport is common because of the large loads applied. Komi (1987) has measured the high forces generated in the human Achilles tendon with the surgical introduction of a calibrated buckle transducer for short periods of time. Forces of up to 4000 N were recorded in the Achilles tendon with toe running. Thus it is not surprising that with repetitive loadings of these magnitudes microfatigue failure could occur with long-distance running, especially in a tendon of small cross-sectional area (Engstrom et al. 1985). Studies by Hasselman et al. (1995) indicate that the proximal tendon of the rectus femoris extends well distal into the muscle belly, and tears at this muscle–tendon junction can appear as mid-muscle belly haematomata.

Wilson and Goodship (1994) have demonstrated a rise of the core temperature of the equine superficial digital flexor tendon under galloping conditions to a mean peak temperature of 43.3°C and a maximum of 45.4°C in one animal. They suggest these core temperatures or central tendon hyperthermia are not survivable by tendon fibroblasts and may explain the central lesions seen in the equine superficial digital flexor tendon and also in the human Achilles tendon.

Spontaneous tendon rupture is uncommon in the young athlete and usually occurs in the older sportsman (for example the long head of the biceps brachii) and then is usually associated with degenerative pathology of the collagen fibrils, although the biochemical detail has not been delineated.

Supraspinatus tendon rupture is common with ageing and may be related to a decreased tendon cross-sectional area and also to increased stiffness of the tissue with increased collagen cross-links. This increased ‘stiffness’ may be further aggravated by matrix changes such as fibrocartilage transformation and frank intratendinous calcification. Recently, Katayose and Magee (2001) have performed an elegant and detailed study of the cross-sectional areas (CSAs) of the normal dominant and non-dominant supraspinatus tendon as a function of age using diagnostic ultrasound with a high level of interobserver reliability. They found that the CSA of the dominant side was significantly larger than the non-dominant side and that the CSA decreased significantly with ageing. They noted that the CSAs of the 70–79-year age group were 82% of those in the 20–29-year age group.

Ligament/tendon repair

There are three main phases of ligament/tendon repair: inflammation, repair (or proliferation) and remodelling (see Figs 4.10–4.14).

Acute inflammation

The gap in the ligament/tendon is filled immediately with erythrocytes and inflammatory cells, especially polymorphonuclear leucocytes. Within 24 h, monocytes and macrophages are the predominant cell and actively engage in phagocytosis of debris and necrotic cells. These are gradually replaced by fibroblasts from either intrinsic or

![Fig. 4.9 Schematic figure of the ultrastructure of the myotendinous junction. Dotted arrow, lamina lucida of the muscular basement membrane; solid arrow, lamina densa of the muscular basement membrane; M, muscle cell; P, muscle cell processes; T, tendon collagen fibrils. (From Kannus et al. 1991, with permission.)](image-url)
extrinsic sources and the initial deposition of the type III collagen scar is commenced. At this stage collagen concentration may be normal or slightly decreased but the total mass of ligament collagen scar is increased. Glycosaminoglycan (GAG) content, water, fibronectin and DNA content are increased (Fig. 4.11).

**Proliferation**

Fibroblasts predominate in this phase. Water content remains increased and collagen content increases and peaks during this phase (3–6 weeks). Type I collagen now begins to predominate and GAG concentration remains high. The increasing amount of scar collagen and reducible cross-link profile has been correlated with the increasing tensile strength of the ligament matrix. Recent quantitative collagen fibril orientation studies indicate that early mobilization of a ligament at this stage (within the first 3 weeks) may be detrimental to collagen orientation. After this time there is experimental evidence that mobilization increases the tensile strength of the repair and probably enhances this phase and the next phase of remodelling and maturation.
With this basic biological knowledge there is now a rationale for the use of early controlled mobilization of patients with ligament/tendon trauma. The use of a limited motion cast with an adjustable double-action hinge for the knee joint is now clinically accepted and enhances more rapid repair and remodelling as well as preserving quadriceps muscle bulk. Patients are now usually mobilized in a limited motion cast for 3–6 weeks rather than the previously empirical time of 6 weeks. An initial range of motion of 20–60° of flexion is used because this places minimal loading on all knee ligaments (Helbing & Burri 1977). The beneficial effects of functional tendon casting versus rigid casting has also been recently demonstrated in a rabbit Achilles tenotomy model.


Fig. 4.12 Ligament repair: (a) with and (b) without RICE (rest, ice, compression and elevation) management and early mobilization.
that demonstrated a 60% increase in tendon collagen and a 20% increase in maximum load and maximum stress compared with control rigid casts when measured 15 days’ post-tenotomy (Stehno-Bittel et al. 1998). Such basic science studies have led to the earlier mobilization of human Achilles tendons after surgical repair (Mandelbaum et al. 1995; Aoki et al. 1998; Speck & Klaue 1998).

Remodelling and maturation
(6 weeks to 12 months)

During this phase there is a decrease in cell numbers and hence a decreased collagen and GAG synthesis. Water content returns to normal and collagen concentration returns to slightly below normal, but total collagen content remains slightly increased. With further remodelling there is a trend for scar parameters to return to normal but the matrix in the ligament scar region continues to mature slowly over months or even years. Scar collagen matrix and adjacent normal matrix may actually shorten the repair region, perhaps by interaction of ligament/tendon myofibroblasts with their surrounding collagen matrix (Danhers et al. 1986). Collagen fibril alignment in the longitudinal axis of the ligament occurs even though small-diameter fibrils are involved (Figs 4.14 and 4.15).

Fig. 4.13 Muscle repair at the muscle–tendon junction with and without RICE (rest, ice, compression and elevation) management and early mobilization.
Occasionally calcium apatite crystals will be deposited in the damaged tissues and the classic site for this to occur is in the rotator cuff supraspinatus tendoperiosteal attachment to the greater tubercle of the humerus.

Achilles tendon and infrapatellar tendon injuries, especially partial tears, present a dilemma for the clinician in that they are often intractable to reasonable short-term management although Stanish et al. (1986) claim good clinical results from graded eccentric loading regimes for patellar tendonitis. The author has examined Achilles tendon biopsies of patients with chronic localized tears and more generalized thickened, tender, chronic Achilles tendons. The feature which characterized the pathology ultrastructurally was the persistence of small-diameter collagen fibrils < 100 nm diameter. The large fibrils of the original tendon do not appear to be replaced in the mature adult in either a repairing tendon or ligament (Fig. 4.15) (see below); increasing the collagen fibril diameters (and their alignment) is needed to enhance the repair scar tensile strength.

ACL injuries appear to be unique in that the chondrocyte-like cells in this special ligament apparently have a limited capacity to proliferate and synthesize a new collagen matrix and hence

---

**Fig. 4.14** The effect of collagen repair with identical fibrils but different geometry and the corresponding load–strain response to tensile testing. (Adapted from Viidik 1980.)

---

**Fig. 4.15**

(a) Number of fibrils versus fibril diameter in patients with chronic Achilles tendonitis; (b) expressed as per cent area occupied for each diameter group versus diameter. The preponderance of small-diameter fibrils in ‘repairing’ chronic Achilles tendonitis should be noted. The large-diameter normal fibrils > 100 nm are not replaced.
repair appears to be limited. Collagenase release may also affect the effectiveness of the repair process (Amiel et al. 1989). However Ng et al. (1996a, 1996b) have completed a long-term biomechanical study of the repair of the goat ACL after hemisection injury showing that, contrary to the current orthopaedic dogma, the ACL can heal even after partial tears. To test the healing capacity of the partially torn ACL, ligaments were tested at 12, 24 and 52 weeks and 3 years after surgery (Fig. 4.16). As early as 12 weeks after surgery, a translucent fibrous tissue covered the transected posterolateral bundle. A comparison of anteroposterior laxity between the right and left knees measured at 45 and 90° of flexion showed no significant difference at each time period. Results of Instron testing of the posterolateral bundle revealed that normalized changes in load–relaxation and Young’s modulus were also not significantly different at each time period, but the ultimate tensile strength and stiffness at 3 years were significantly higher than at 12 weeks \( P < 0.05 \). Failure started at the repair site for the 12-week group, but at 24 and 52 weeks the failure occurred through the ligament. At 3 years, the posterolateral bundle specimens failed by bony avulsion, indicating the repaired tissue was not the weakest link of the bone–ligament–bone complex.

This study showed that under the favourable biological conditions of the experiment, partial ACL injuries in the goat are capable of adequate mechanical repair. What is more important, the high ultimate tensile strength and stiffness of the 3-year repaired tissue indicate that full structural repair of such an artificial transection injury is possible (Figs 4.16–4.18).

**Clinical and ultrastructural observations of Achilles tendon injuries**

In this section I will attempt to relate the three phases of healing of soft tissues with that seen in Achilles tendon injuries. Achilles tendon injuries can be classified as described above for ligament injuries, i.e. grades 1–3 with the latter being complete rupture. Several patient histories will be used to illustrate these three phases of healing and attempted repair.
Grade 3: complete rupture

Australian Rules football rover, aged 26, powering off to avoid an opponent. Clinical signs were a palpable defect in the Achilles tendon as well as a positive Thompson’s sign. A biopsy was taken during surgical repair.

With this disastrous injury there is both collagen bundle failure as well as vascular disruption and hence bleeding is a feature of these injuries when seen at early surgical repair. This trauma to the tendon initiates the acute inflammatory phase and hence microscopically there is massive red cell extravasation, fibrin clot formation and collagen fibril disruption. The damaged tendon becomes oedematous and polymorphonuclear monocytes migrate into this area and release their lysosomal contents as well as actively phagocytosing cellular and other debris. Macrophages also move into the rupture site and commence phagocytosis of damaged cells and tissue. This phase lasts 0–72 h or so and is followed by the repair phase (Fig. 4.19).

Grade 2: tear

Australian Rules football ruckman with 6 months’ painful thickened Achilles tendon. The operation involved the excision of paratenon and tendon incision to remove damaged, haemorrhagic and
repair of tendons and ligaments

necrotic regions. A biopsy was taken for light and electron microscopy.

Light microscopy demonstrated a thickened paratenon and an oedematous thickened tendon (Fig. 4.20). Ultrastructurally many fibroblasts had dilated rough endoplasmic reticulum indicative of increased collagen synthesis. Apart from the many free red cells, the other feature was the prevalence of many small-diameter collagen fibrils not aligned or closely packed (18–20 nm diameter) in amongst the older, larger, pre-existing fibrils ranging from 80 to 150 nm diameter. Polymorphonuclear monocytes and macrophages were not common at this stage, reflecting perhaps the slowness of repair in this unique tissue.

The biopsy demonstrated the features of the repair phase that follows the acute inflammatory phase and lasts 72 h to 4–6 weeks; but in this patient the repair phase had been perpetuated because of continued activity by the athlete. In other patients where there is less florid tendonitis and the tendon is clinically painful but not obviously enlarged, discrete areas of increased cellularity are seen in the tendon in the form of free red cells, cell debris and viable fibroblasts surrounded by both large- and many small-diameter collagen fibres. These areas are almost ‘walled-off’ from the densely packed collagen of the rest of the tendon by a fibrin precipitate. These discrete areas are probably due to collagen fibre ruptures or microtears corresponding to the early part of the second region of the load–deformation curve. There is as yet no evidence that the use of massage or deep friction enhances this repair phase (Walker 1984).

Grade 1: injury

Runner, aged 25, with a painful tender lump in the Achilles tendon for 18 months! A biopsy was obtained at open operation.

The lump at operation was firmer than the rest of the tendon and was slightly darker in colour. On light microscopy of the biopsy taken from the nodule, the changes in the collagen bundles were very subtle. There was less regular ‘crimping’ of the collagen bundles and they were not as tightly packed. However, ultrastructurally the cause of this less regular collagen crimp was obvious in that between the large-diameter fibrils were many small-diameter fibrils that were less well orientated longitudinally (Fig. 4.15). These tendons were interpreted as being in the remodelling phase as there were no increased fibroblast numbers in the nodules and no inflammatory cells were observed (Fig. 4.15).

The mechanism of acute complete rupture in young athletes ‘powering-off’ during sprinting indicates that the gastrocnemius–soleus complex can generate sufficient force to rupture the tendon. However, tendon strength usually exceeds that of its muscle by a factor of two and hence rupture is unusual. The mechanisms involved in partial
grade 1 and 2 tears in the Achilles tendon are not as obvious. Viidik (1973) has shown that rat tendons \textit{in vitro} undergo increasing deformation or ‘plasticity’ if cycled to loads of less than one-tenth of its failure load, and that the strain or deformation occurs well before region two or the linear part of the load–strain curve begins (Fig. 4.21). Similar observations have been made \textit{in vivo} and \textit{in vitro} for rat knee joint ligaments by Weisman \textit{et al} (1980). It is possible that in distance runners a similar fatigue plasticity and elongation occurs in the tendon and this causes the microruptures and the thickened repair nodules already described.

The notion that some running athletes may not have an Achilles tendon of sufficient CSA to sustain the repetitive tendon loading of distance without injury has been investigated by Engstrom \textit{et al} (1985). In an elegant study they used ultrasound to measure the CSA of the human Achilles tendon \textit{in vivo} and validated this technique as a reliable method to measure the Achilles tendon CSA using cadaver Achilles tendons. Two groups of distance athletes with and without grade 1 Achilles tendonitis who were age, weight and distance matched had their Achilles tendon CSA measured using the previously validated ultrasound technique. They demonstrated that athletes with grade 1-type Achilles tendonitis had about a 30% decrease in the CSAs of their Achilles tendons (Fig. 4.22). This indicates that a major mechanism in this type of common injury may simply be fatigue creep failure of Achilles tendon collagen as shown in Fig. 4.21. Komi \textit{et al} (1987) have developed an \textit{in vivo} buckle transducer which they located around the Achilles tendon in a number of subjects. Direct force measurements were made on several subjects who were involved in slow walking, sprinting, jumping and hopping after calibration of the transducer. During running and jumping forces close to the previous estimated ultimate tensile strength of the tendon were recorded, indicating that fatigue creep in a small cross-sectional tendon is a possible mechanism of injury without the need to invoke other lower limb biomechanical pathology as has been suggested by Clement \textit{et al} (1984) and Williams (1986).
Use of physical modalities in attempts to increase collagen deposition for enhancing the tensile strength of soft tissue repair: electrical, ultrasound and laserphoto stimulation

The original in vitro work of Fitton-Jackson and Bassett (1980), in which fibroblasts were stimulated to increase collagen synthesis under the influence of pulsed magnetic fields, has been in part confirmed by other workers. However, the gains in increased ligament strength have been meagre compared to the dramatic 40% or so reduction in healing times for fresh fractures of the tibia and radius with the use of low-intensity ultrasound applied to fracture healing as originally described in the human by Heckman et al. (1994), and recently reviewed by Rubin et al. (2001).

In fracture repair, ultrasound was demonstrated with substantive data to enhance the three basic stages of the healing process described above. Such observations have been also demonstrated in part in soft tissue repair, but not with the same large increases in matrix deposition (collagen) as for bone deposition. Frank et al. (1983a, 1983b) demonstrated accelerated early healing of a rabbit medial collateral ligament (MCL) to which a solid core electromagnet energized the tissues with a square wave of unidirectional current 7 h per day for up to 6 weeks postinjury. An increased tensile strength of almost double the control MCL was observed at 21 days, but at 42 days there was no difference between the treated and control ligaments.

Enwemeka (1989) applied therapeutic ultrasound daily for 5 min for nine treatments to tenotomized rabbit Achilles tendons, and noted significant increases in tensile strength and energy absorption to failure in the early healing phase (10 days’ post-tenotomy).

Kesava Reddy et al. (1998) applied a combined laser treatment and electrical stimulation to the repairing rabbit Achilles tendon after complete tenotomy over a short 5-day period. No statistical difference was found between the experimental and the control tenotomies in terms of mechanical performance although there was a 32% increase in collagen deposition at the repair site. Previous studies using low-intensity laser did moderately increase the biomechanical load-bearing capacity of healing rabbit Achilles tendons after 14 days of treatment (Enwemeka et al. 1990).

Clearly, more long-term studies need to be done in this area to attempt to understand the biological mechanisms involved and also to determine the long-term therapeutic value of such treatments.

Effects of immobilization on the capsule and synovial joints

Akeson et al. (1987) reviewed their work and others’ work on this important topic. There is articular cartilage atrophy with proteoglycan loss and an associated fibrofatty connective tissue that is synovial derived, which adheres to the articular cartilage. Ligament insertion sites are weakened and the ligament itself has increased compliance with reduced load to failure due to loss of collagen mass, which may occur with only 8–12 weeks’ immobilization but may take up to 1 year to recover after mobilization (Amiel et al. 1983). Capsular changes include loss of water due to loss of GAGs and hyaluronic acid, leading to joint stiffness. Clearly joint immobilization is to be avoided if possible to prevent the above changes from occurring which take many months to recover (Fig. 4.23). Similarly Steno-Bittel et al. (1998) have demonstrated that functional casting of rabbits after an Achilles tendon complete transection and surgical repair in a short-term experiment (15 days); increased collagen deposition (60%) and increased maximum tensile strength (20%) and maximum stress (21%) in the repairing tendon compared to complete Achilles tenotomies and suture repair with classic immobilization of the lower limb with non-functional load bearing.

Thus the thrust in recent years has been to minimize immobilization times and to avoid soft and hard tissue immobilization if practically possible. With wounds to soft tissues and hard tissue injuries, such as fractures to bone, it may be necessary to immobilize these tissues either
externally (or internally with surgery) until healing is well underway. This is so these tissues will have developed enough intrinsic mechanical strength through neocollagen fibril deposition so as not to be disrupted by early mobilization and to allow the remodelling phase of the repair process to be enhanced (Vailas et al. 1981; see below). Both tensile-bearing soft tissues such as tendon and ligament require regular minimal loading to maintain their intrinsic mechanical tensile strength; they are similar to bone and articular cartilage in this respect (Caterson & Lowther 1978; Rubin & Lanyon 1987).

Effect of immobilization and mobilization on ligament tensile strength

Effect of immobilization

The most important study in this context is that by Noyes et al. (1974) using the ACL in monkeys. They demonstrated clearly that 8 weeks of lower limb cast immobilization led to a substantial loss of ligament tensile strength which took 9 months to recover from, even with a reconditioning programme. The predominant mode of failure in these experiments was ligament failure as opposed to avulsion fracture failure. The latter mode of failure was more predominant in the immobilized group due to resorption of Haversian bone at the ligament attachment; but after 5 months of reconditioning femoral avulsion fractures did not occur. It should be noted that there was no surgery, just simple immobilization.

Another important study with direct implications for clinical practice is that of Amiel et al. (1985). Although this work was done in rabbits, in the context of the previous work by Noyes, it is very relevant. The timeframes for the changes in ligament tensile strength are similar and hence are applicable to clinical rehabilitation. These investigators were able to show that 12 weeks of immobilization of the MCL in the growing rabbit led to profound atrophy, such that there was a decrease of approximately 30% in collagen mass as a result of increased collagen degradation. Remarkably most of this atrophy occurred during weeks 9–12 of immobilization. Again, there was no trauma or surgery to the MCL in this study.

Hence it appears from the above two studies that prolonged immobilization, i.e. 6–12 weeks without trauma, can itself lead to profound atrophy of both the collateral and cruciate ligaments of the knee joint and recovery may require at least many months or even a year. This timeframe must be kept in mind when managing patients after prolonged knee immobilization and advising them when they can best return to full competitive sport.

Amiel et al. (1985) have also shown that there is a close relationship between joint stiffness induced by immobilization and a decrease of total GAG in the periarticular connective tissues. They demonstrated alleviation of this joint stiffness in a rabbit model using intra-articular hyaluronate.

Apart from the original, classic work of Noyes et al. (1974) in which the effects of immobilization on the ACL of the primate were examined, there have been few investigations of the cause of the decreased strength and elastic stiffness of the ACL in response to immobilization. Tipton et al. (1970) reported that collagen fibre bundles (viewed with light microscopy) were decreased in number and size in immobilized dogs, suggesting
that this was the cause of the decreased CSA seen in immobilized rabbit MCLs (Woo et al. 1987a). The explanation for the decreased strength and elastic stiffness in these immobilized ligaments may be found at the collagen fibril level. Binkley and Peat (1986) showed a decrease in the number of small-diameter fibrils after 6 weeks of immobilization in the rat MCL.

**Effect of mobilization (exercise)**

There have been a large number of studies investigating this area. The literature has been reviewed by Tipton et al. (1986) Tipton and Vailas (1989), Butler et al. (1979) and Parker and Larsen (1981). The results in experimental animals generally indicate an increase in bone–ligament–bone preparation strength as a response to endurance-type exercise. However, no change in ligament or tendon strength has been recorded by some workers and this may reflect different exercise regimes, methods of testing as well as species differences.

The observations by Tipton et al. (1970, 1975), Cabaud et al. (1980), Parker and Larsen (1981) and others that ligament strength is dependent on physical activity prompted an ultrastructural investigation as to the mechanism of this increase in tensile strength. Increased collagen content was found in the ligaments of exercised dogs and this correlated with increased CSA and larger fibre bundles. This accounts for the increased ligament tensile strength but whether this increased collagen was due to the deposition of collagen on existing fibres or due to the synthesis of new fibres was not investigated. Larsen and Parker (1982) had already shown that with a 4-week intensive exercise programme in young male Wistar rats, both the ACLs and PCLs showed a significant strength increase ($P < 0.05$).

Oakes et al. (1981) and Oakes (1988) quantified the collagen fibril populations in young rat ACLs and PCLs subjected to an intensive 1-month alternating treadmill and swimming exercise programme. This study was performed in an attempt to explain the increased tensile strength found in these ligaments with the intensive endurance exercise programme and to determine if this could be explained at the level of the collagen fibril, which is the fundamental tensile unit of ligament.

Five 30-day-old pubescent rats were placed on a progressive 4-week exercise programme of alternating days of swimming and treadmill running. At the conclusion of the exercise programme the rats were running 60–80 min at 26 m·min$^{-1}$ on a 10% treadmill gradient and on alternate days swimming 60 min with a 3% body weight attached to their tails. Five caged rats of similar age and commencing body weights were controls. Analysis of ultrathin transverse sections cut through collagen fibrils of the exercised ACLs revealed: (i) a larger number of fibrils per unit area examined (29% increase, $P < 0.05$) compared with the non-exercised caged control ACLs; (ii) a fall in mean fibril diameter from 9.66 ± 0.3 nm in the control ACLs to 8.30 ± 0.3 nm in the exercised ACLs ($P < 0.05$); and (iii) as a consequence of (i) and (ii) the major CSA of collagen fibrils was found in the 11.25 nm diameter group in the exercised ACLs and in the 15 nm diameter group in the control ACLs. However, total collagen fibril cross-section per unit area examined was approximately the same in both the exercised and the non-exercised control ACLs. Similar changes occurred in the exercised and control PCLs. These results are shown in Figs 4.24–26. In the exercised PCL, collagen per microgram of DNA was almost double that of the control suggesting that the PCL was more loaded with this exercise regime than the ACL. The conclusion from this study is that ACL and PCL fibroblasts deposit tropocollagen as smaller diameter fibrils when subjected to an intense 1 month’s intermittent loading (exercise) rather than the expected accretion and increase in size of the pre-existing larger diameter collagen fibrils.

Very similar ultrastructural observations have been made for collagen fibrils of exercised mice flexor tendons (Michna 1984).

The mechanism of the change to a smaller diameter collagen fibril population is of interest and may be related to a change in the type of proteoglycans synthesized by ligament fibroblasts in response to the intermittent loading of exercise.
It is well recognized since the original work of Toole and Lowther (1968) that GAGs have an effect on determining collagen fibril size in vitro and this has been confirmed in vivo by Parry et al. (1982). Merrilees and Flint (1980) demonstrated a change in collagen fibril diameters between the compression and tension regions of the flexor digitorum profundus tendon as it turns 90° around the talus. Amiel et al. (1984) have also shown that rabbit cruciate ligaments have more GAGs than the patellar tendon and hence it is likely that GAGs also play an important role in determining collagen fibril populations in cruciate ligaments (see below for further discussion).

**Surgically Repaired Ligaments**

Tipton et al. (1970) demonstrated a significant increase in the strength of surgically repaired MCLs of dogs after treadmill exercise training for 6 weeks, after 6 weeks’ cast immobilization.
However, they emphasized that at 12 weeks’ postsurgery (6 weeks of immobilization and 6 weeks of exercise training) the repair was only approximately 60% that of the normal dogs and results suggested that at least 15–18 weeks of exercise training may be required before a return to ‘normal’ tensile strength is achieved. Similar observations have been made by Piper and Whiteside (1980), using the MCL of dogs. They observed that mobilized MCL repairs were stronger and stretched out less, i.e. less valgus laxity, than MCL repairs managed by casting and delayed mobilization. This conclusion is supported by the work of Woo et al. (1987a, 1987b).

Some insight into the biological mechanisms involved in the repair response with exercise has come from the elegant work of Vailas et al. (1981). By using 3H-proline pulse labelling to measure collagen synthesis in rat MCL surgical repairs and coupling this with DNA analyses and tensile testing of repaired ligaments subjected to exercise and non-exercise regimes, they were able to show that treadmill running exercise commencing 2 weeks after surgical repair enhanced the repair and remodelling phase by inducing a more rapid return of cellularity, collagen synthesis and ligament tensile strength to within normal limits. Further, Woo et al. (1987b) have demonstrated almost complete return (98%) of structural properties of the transected canine femoral–medial collateral ligament–tibial (FMT) complex at 12 weeks’ post-transection without immobilization. Canines immobilized for 6 weeks with the FMT complex tested at 12 weeks, had mean loads to failure of 54% that of the controls. However, the tensile strength of the MCL was only 62% of controls at 48 weeks. This apparent paradox in the non-immobilized dogs was explained by the approximate doubling in CSA of the healing MCL (and hence increased collagen deposition) during the early phases of healing. This repair collagen was most probably small-diameter fibrils and would account for the poor strain performance of the MCL scar collagen and would be similar to the lower curve on Fig. 4.23.

Woo et al. (1987a) examined the effects of prolonged immobilization and then mobilization on the rabbit MCL without any surgical inter-vention. Both the structural properties of the FMT complex and the material properties of the MCL were examined. After immobilization, there were significant reductions in the ultimate load and energy-absorbing capabilities of the bone–ligament–bone complex. The MCL became less stiff with immobilization and the femoral and tibial insertion sites showed increased osteoclastic activity, bone resorption and disruption of the normal bone attachment to the MCL. With mobilization, the ultimate load and energy-absorbing capabilities improved but did not return to normal. The stress–strain characteristics of the MCL returned to normal, indicating that the material properties of the collagen in rabbit MCL return relatively quickly after remobilization but that the ligament–bone junction strength return to normal may take many months (see Fig. 4.23).

The detailed biological cellular mechanisms involved in this enhancement and remodelling of the repair are not understood but may involve prostaglandin and cAMP synthesis by fibroblasts subjected to repeated mechanical deformation by exercise.

Amiel et al. (1987) have shown that maximal collagen deposition and turnover occurs during the first 3–6 weeks postinjury in the rabbit. Chaudhuri et al. (1987) used a Fourier domain directional filtering technique to quantify collagen fibril orientation in repairing ligaments. Results indicated that ligament collagen fibril reorientation does occur in the longitudinal axis of the ligament during remodelling. MacFarlane et al. (1989) and Frank et al. (1991) have further shown that collagen remodelling of the repairing rabbit MCL appears to be encouraged by early immobilization but after 3 weeks collagen alignment and remodelling appears to be favoured by mobilization.

We have recently completed a study on the repair of the goat ACL. It has almost become an axiom amongst orthopaedic surgeons that the ACL does not repair after injury, even partial injury. This seems to be the case with complete ACL rupture (Grontvedt et al. 1996) if there is no continuity in either the synovial membrane ACL sleeve or its contained collagen fibre bundles. However, the repair of partial tears of the ACL
(grade 1 and 2) has not been examined carefully in the long term in large animals. To test the healing of the partially torn ACL, we transected the posterolateral bundle of the ACL in 11 adult female goats and tested the ligaments at 12, 24 and 52 weeks and 3 years after surgery as previously described on page 69. Translucent repair tissue was seen filling in the ‘wound’ region as early as 12 weeks post-hemisection. There was also no difference in the antero-posterior laxity between the hemisected and normal control knees at each time period examined. Load-relaxation and Young’s modulus also were not significantly different at each time period; but surprisingly the hemisected-ACL ultimate tensile strength and stiffness at 3 years were significantly higher than at 12 weeks \( (p < 0.05) \). Also surprising was the fact that at 3 years, the ACL specimens all failed with bone avulsion indicating the repair tissue was very strong and not the weakest link of the bone–ligament–bone complex as one might have expected (see Figs 4.16 and 4.17).

This study emphasizes that under favourable conditions, partial ACL injuries or tears are capable of an adequate tensile strength repair. Perhaps more importantly as far as athletes are concerned the high ultimate tensile strength and stiffness of the 3-year repaired tissue indicates that full structural repair of such an artificial hemitranssection injury model is possible (Ng et al. 1996). This work suggests such partial human ACL injuries may be adequately repaired in athletes, (if diagnosed early with MRI examination) which is contrary to current orthopaedic opinion.

With this basic biological knowledge there is now a rationale for the use of early controlled mobilization of patients with ligament trauma. The use of a limited motion cast with an adjustable double-action hinge for the knee joint is now accepted in clinical practice and enhances more rapid repair and remodelling as well as preserving quadriceps muscle bulk and strength. Patients are now usually mobilized early in a limited motion cast for a minimum of 3 weeks with a range of motion from 20 to 60° (Helbing & Burri 1977) rather than the previously empirical time of 6 weeks’ immobilization.

Collagen fibril populations in human knee ligaments and grafts

Human ACL autograft quantitative collagen fibril studies

In order to gain some biological insight into collagen repair and remodelling mechanisms within human cruciate ligament grafts, biopsies were obtained from autogenous ACL grafts from patients subsequently requiring arthroscopic

Fig. 4.27 Transverse sections through collagen fibrils of: (a) normal young adult patellar tendon (mean of six biopsies); (b) normal young adult ACL (mean of six biopsies); and (c) Jones’ free graft (mean of nine biopsies). (Magnification × 34 100.) The insets show: on the left, the number of fibrils versus diameter; and on the right, per cent area occupied/diameter group. The preponderance of small-diameter fibrils in the graft (c), and large fibrils in the patellar tendon (a), which are not seen in the ‘normal’ ACL (b), should be noted.
intervention because of stiffness, meniscal and/or articular cartilage problems or removal of prominent staples used for fixation. Most of the ACL grafts were from the central one-third of the patellar tendon as a free graft (n = 39), or in some left attached distally (n = 8) and in others the hamstrings (n = 9, graft age 10 months to 6 years) or the iliotibial tract was used (n = 15, graft age 10 months to 6 years). These biopsies represented approximately 20% of the total free grafts performed over the 3 years of this study. The clinical ACL stability of the biopsy group differed little from the remainder. All had a grade 2–3 pivot shift (jerk) preoperatively (10–15 mm anterior drawer neutral), eliminated postoperatively in 87% of patients (0.5 mm). Subsequent clinical review at 3 years showed an increase in the anterior drawer with a return in 20% of a grade 1 pivot shift.

A total of 39 biopsies have been quantitatively analysed for collagen fibril diameter populations in patients aged 19–42 years. These data were compared with collagen fibril populations obtained from biopsies of cadaver ACLs (n = 5) and also biopsies of ACLs from young (<30 years, n = 5) people who had not sustained a recent ACL injury. Biopsies were also obtained from normal patellar tendons at operation (n = 7) (Figs 4.27 and 4.28). Eighteen ACL graft biopsies have been obtained from other surgeons, both nationally and internationally (Oakes et al. 1991; Oakes 1993).

The results from the collagen fibril diameter morphometric analysis in all ACL grafts clearly indicated a predominance of small-diameter collagen fibrils (Figs 4.28–4.30). Absence of a ‘regular crimping’ of collagen fibrils was observed by both light and electron microscopy, as was a less ordered parallel arrangement of fibrils. In most biopsies, capillaries were present and most fibroblasts appeared viable.

**Collagen typing of normal human ACL and ACL grafts**

Recent biochemical analyses of human patellar tendon autografts in situ for 2–10 years indicates a large amount of type III as well as type V collagen. This confirmed our suspicion that a large
Fig. 4.29  Summary histograms of all human anterior cruciate ligament (ACL) patellar tendon (PT) autografts ($n = 39$) versus time postgrafting compared with normal young ACL ($n = 10$) and normal young PT ($n = 7$) expressed as per cent area per diameter group. The presence of large fibrils (>100 nm) in the older (5-9-year-old) grafts; arrow 1) and the rapid loss of <100 nm fibrils in the 0-6-month postgraft group (arrow 2), which are present in the donor PT (to the left of arrow 2) should be noted. (From Oakes 1993, with permission from Raven Press.)

Fig. 4.30  Summary histograms of all human anterior cruciate ligament (ACL) allografts versus time postgrafting compared with normal young ACL, normal patellar tendon (PT), reconstituted tibialis anterior (TA) and Achilles tendons expressed as per cent area per diameter group. The early preponderance of the small fibrils at 3 months postgraft (arrow 1) and their persistence even at 96 months postgraft in the majority of biopsies examined should be noted. Some large fibrils are present (arrow 2) at 96 months, but the majority of the large fibrils present in the donor tissues are removed (to the right of arrow 3). (From Oakes 1993, with permission from Raven Press.)
amount of collagen in these remodelled grafts at this age may be type III and not type I as is normally found in the patellar tendon and adult ACL (Deacon et al. 1991).

We have examined and quantified the types of collagen in the normal ACL and in ACL grafts using quantitative sodium dodecyl sulphate gel densitometry of cyanogen bromide peptides derived from the tissue in question (Chan & Cole 1984). Tissue was obtained from: 10 acute ruptured ACLs (< 1 week old), 10 normal ACLs and nine autogenous patellar tendon grafts (age 3 months to 2 years).

The normal ACL contained a mean type I collagen content of 71.13 ± 9.77 (SD) and type III content of 28.1 ± 10.18. The acute ACL ruptures had similar type I and III collagen contents as did the patellar tendon grafts (Table 4.1).

The high content of type III collagen in the normal human ACL (28.1%) is surprising and was similar to that found in the patellar tendon autografts and the acute ACL ruptures. This is much higher than that reported by Amiel et al. (1984) for normal rabbit ACL. They suggest the high type III content may reflect a wide variety of force vectors that the rabbit ACL is subjected to. These new data may indicate unrecognized or documented previous injury to the ‘normal’ ACL tissue (Oakes 1993).

Before discussing the biopsy data it is of interest to compare the collagen profiles for the patellar tendon with the normal ACL. It can be seen that the profiles are different in that the distribution in the patellar tendon is skewed to the right with a small number of large fibrils not present in the normal ACL (Figs 4.27–4.30). Elegant work by Butler et al. (1985) has shown that the patellar tendon is significantly stronger than the human ACL, PCL and lateral collateral ligament from the same knee in terms of maximum stress, linear modulus and energy density to maximum strength. The larger fibrils observed in the patellar tendon are not found in the ACL and are an obvious explanation for the stronger biomechanical tensile properties of the normal patellar tendon.

The biopsies from the grafts were obtained from patients with a good to fair rating in terms of a moderate anterior drawer (0–5 mm) and correction of the pivot shift, but both these tests of ACL integrity showed an increasing laxity of the ACL at the 3-year clinical review. The length of time the grafts were in vivo prior to biopsy varied from 6 months to 6 years. The collagen fibril population did not alter that much for the older grafts (namely > 3 years), which is not what was hoped for or expected but is in keeping with the observation clinically that the ACL grafts ‘stretched out’ postoperatively.

The most striking feature of all the biopsies from the autografts, irrespective of whether they were ‘free grafts’, Jones’ grafts, fascia lata, hamstring grafts, and independent of the surgeon, was the invariable prevalence of small-diameter fibrils in amongst a few larger fibrils that probably were the original large-diameter patellar fibrils. The packing of the small fibrils in the

### Table 4.1 Human ACL collagen typing.

<table>
<thead>
<tr>
<th>Type</th>
<th>Normal ACL (N = 10, age 16–40)</th>
<th>Acute ACL rupture (n = 11, age 16–30, &lt; 1-week-old)</th>
<th>Jones’ free ACL grafts (n = 9, age 19–27, graft age 9–2yrs)</th>
<th>Iliotibial band ACL grafts (n = 3, age 25–35, graft age 4–10yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (+/− SD)</td>
<td>71.13 +/− 9.77</td>
<td>65.51 +/− 7.80</td>
<td>70.20 +/− 6.75</td>
<td>60.27 +/− 13.89</td>
</tr>
<tr>
<td>Type III (+/− SD)</td>
<td>28.1 +/− 10.18</td>
<td>33.77 +/− 7.86</td>
<td>28.48 +/− 7.02</td>
<td>39.03 +/− 13.90</td>
</tr>
</tbody>
</table>
grafts was not as tight as is usually observed in the normal patellar tendon (compare Figs 4.27a, c and 4.29, arrows 1 and 2).

It appears from the quantitative collagen fibril observations in this study that the large-diameter fibrils of the original graft are removed and almost entirely replaced by smaller, less well-packed and less well-orientated fibrils than the larger diameter fibrils found in the normal patellar tendon. The smaller diameter fibrils are probably recently synthesized because they are of smaller diameter than those found in the original patellar tendon. Hence the entire collagen matrix of the original patellar tendon is remodelled and replaced with newly synthesized small-diameter collagen fibrils.

Is gentle mechanical loading in the ACL grafts an important stimulus to fibroblast proliferation and collagen deposition? Inadequate mechanical stimulus may occur, especially if grafts are non-isometric and are ‘stretched out’ by the patient before they have adequate tensile strength. A lax ACL graft may not induce sufficient mechanical loading on graft fibroblasts to alter the GAG or decorin/collagen biosynthesis ratios to favour large-diameter fibril formation. Certainly in this study there was ACL graft laxity which increased postoperatively. This would lend credence to the above notion. However, use of continuous passive motion in grafted primates does not increase the strength of grafts.

Another more likely possibility is that the replacement fibroblasts in the ACL grafts are derived from stem cells from the synovium (and synovial perivascular cells) which are known to synthesize hyaluronate, which in turn favours small-diameter fibril formation (Parry et al. 1982). The strong correlation of small-diameter fibrils with a lower tensile strength has been observed by Parry et al. (1978) and Shadwick (1990), and the observations in this study would confirm this. Our observations also correlate with the observations of Clancy et al. (1981) and Arnoczky et al. (1986). The observations by Amiel et al. (1989) indicate that collagenase may play a role in the remodelling of ACL tears/grafts.

The conclusion from these ultrastructural observations is that the predominance of small-diameter collagen fibrils (< 7.5 nm) and their poor packing and alignment in all the ACL grafts, irrespective of the type of graft (auto- or allograft), their age and the surgeon, may explain the clinical and experimental evidence of a decreased tensile strength in such grafts compared with normal ACL. It appears that in the adult, the replacement fibroblasts in the remodelled ACL graft cannot reform the large-diameter, regularly crimped and tightly packed fibrils seen in the normal ACL, even after 6 years, which was the oldest graft analysed.

The origin of the replacement fibroblasts which remodel the ACL grafts is not known at the moment. It is the author’s hunch that they will not come from the actual graft itself although some of these cells may survive due to diffusion. However, the bulk of the stem cells involved in the remodelling process are probably derived from the surrounding synovium and its vasculature.

Beynon et al. (1997) mechanically tested a human patellar tendon ACL graft 8 months post-surgery. The stiffness and the ultimate failure load approached that of normal after 8 months of healing and graft remodelling. The anteroposterior (AP) laxity was 1.85 and 1.26 that of the normal knee at 10 and 60° of knee flexion, respectively. This increase in AP laxity (6.3 mm greater than normal) was substantially greater than the 2 mm right-to-left variation in AP laxity seen in subjects with healthy knees. This was also reflected in greater strain values when compared with the normal ACL. Also, energy absorbed to failure was only 53% of the normal ACL. These values of increased strain, increased AP laxity and low energy to failure indicate the poor quality of the remodelled collagen matrix of this ACL and is similar to that of the scar seen in the repairing dog (Woo et al. 1987b) and rabbit MCL (Frank et al. 1999; see below).

Human ACL allografts

Recent further studies of biopsies obtained from ACL human allografts utilizing fresh frozen Achilles or tibialis anterior tendons, ranging in age from 3 to 54 months, indicated a similar

Human ACL allograft specimens were studied as above for the autografts with quantitative collagen fibril analyses (Oakes 1993), and were compared with ACL autografts. The allograft specimens were procured at the time of second-look arthroscopy from the superficial region of the mid-zone of ACL grafts after synovial clearance. The grafts used for the ACL reconstruction were usually from fresh frozen allogeneic Achilles or tibialis posterior or anterior tendons and were implanted 3–96 months prior to biopsy.

Thirty-eight patients who had undergone allograft ACL replacement and whose AP stability had been adequately restored were randomly selected. The restored stability of the involved knees was carefully confirmed with both Lachman and pivot shift signs and an objective quantitative knee instability testing apparatus. All of these patients were subjected to second-look arthroscopy as a part of the procedure to remove hardware installed for graft fixation. Thirty-five graft biopsies were obtained from this patient group. Their age ranged from 15 to 37 years at the time of reconstruction.

ACL reconstruction was performed using fresh frozen allografts, 8–9 mm in diameter; part of the Achilles tendon, the tibialis anterior or posterior, peroneal or other thick flexor tendons without any bone attached to their ends were used as an ACL substitute. Postoperatively, the knee was immobilized for 2–5 weeks, then full weight bearing allowed at 2–3 months, jogging recommended at 5–6 months, and full activity allowed at 9–12 months.

The normal tissues used were compared with the normal ACL and the ACL allografts.

Normal tissues used for allografts

The reconstituted Achilles tendon demonstrated a large number of fibrils in the 90–140 nm range (40% of total CSA) together with small-diameter fibrils of 30–80 nm. The reconstituted tibialis anterior tendon showed larger diameter fibrils and fewer smaller fibrils than the Achilles tendon.

The large fibrils constituted about 80% of the total fibril CSA. In contrast, the normal ACL had about 85% of its total CSA composed of fibrils < 100 nm, but there were a small number of large fibrils which accounted for about 15% of its total CSA.

Allograft results versus time (Fig. 4.30)

By 3 months postoperatively (n = 2), there was a predominance of small-diameter fibrils which accounted for > 85% of the total CSA of these biopsies with a ‘tail’ of larger fibrils making the fibril distribution bimodal in shape. At 6 months (n = 5), the fibril distribution was now unimodal with most (c. 90%) of the fibril CSA in the < 100 nm diameter group. Fibrils of > 100 nm were obviously fewer than in the 3-month specimens. By 12 months (n = 12) almost all the CSA resided in the < 100 nm diameter fibril group and there was almost complete absence of the large > 100 nm fibrils.

This profile persisted in the 13–96-month-old allografts. However, there were two exceptions, where one 12-month-old and one 54-month-old ACL allograft biopsy specimen had significant numbers of large-diameter fibrils in contrast to the earlier observations. In these two specimens the large-diameter fibrils (> 100 nm) accounted for 40% of the total CSA of the collagen fibrils. However, the large-diameter fibrils in these exceptional grafts were smaller in size and had more irregular surfaces than those in the allografts prior to implantation, suggesting perhaps a ‘collagenase sculpting’ of their exposed surface fibrils.

Parry et al. (1978) were the first to describe the bimodal distribution of the collagen fibrils in adult mature tendon collagen that is subjected to high tensile loads. The reconstituted (treated by freezing and thawing) human allografts (Achilles and tibialis tendons) and the normal ACL in this study are also shown to have a bimodal distribution of small- and large-diameter fibrils similar to that described by Parry and colleagues in a large range of other tissues.

Most ACL allograft biopsies also demonstrated a bimodal distribution of large and small fibrils
similar to the ‘normal reconstituted’ tendons up until about 6 months postimplantation. However, after this time the distribution became more unimodal, such that there was an increasing predominance of small-diameter fibrils with a concomitant and progressive loss of the larger ‘host tendon’ fibrils. It is these larger fibrils in the ‘normal’ tendon that are responsible for a large percentage of the tendon collagen CSA and for the very high tensile strength of these tendons; the tensile strength exhibited is probably due to the high density of intermolecular collagen cross-links. These observations suggest, at the least, that most of the original large-diameter fibrils in the ACL allografts are replaced by newly synthesized smaller diameter collagen fibrils (or undergo disaggregation, see below). The loss of the large-diameter fibrils from 6 months and older allografts and the predominance of the small-diameter fibrils seen in this study is the most likely explanation for the dramatic reduction in tensile strength of ACL allografts, and hence the observed increased AP laxity in a previous animal study by Shino et al. (1984).

These observations parallel those described above for ACL autografts, which also demonstrated within 6 months postimplantation the loss of large-diameter collagen fibrils of patellar tendon origin. It could be concluded, therefore, that the remodelling process has a similar time-frame in both ACL tendon allografts and patellar tendon ACL autografts. It further suggests that similar mechanisms of collagen degradation and neosynthesis of collagen may be occurring by the invading synovial stem cells which repopulate the ACL graft almost immediately after surgery.

A valid criticism of these studies could be that all the biopsies were obtained from the superficial region of the grafts and that this may not be representative of the bulk of the graft collagen. However, in the previous autograft study the biopsies were usually obtained from the middle of the autografts and the observations were no different from those in this allograft study. This suggests that the superficial region from which the biopsies were taken in this study is representative of the graft collagen.

Awaiting the appearance of large fibrils with close packing (perhaps under a Wolff’s law tensile stimulus) as seen in normal ligament and tendon is probably futile. There is no current ultrastructural evidence available that large-diameter fibrils will eventually be formed and become a large proportion of the CSA of long-term ACL grafts or even that the packing of the small-diameter fibrils becomes closer within these grafts. If denser collagen fibril packing could be achieved, this might possibly enhance the collagen fibril CSA per unit area, which in turn might possibly increase the tensile strength of the graft by increased interfibril interactions as already discussed. The reason denser packing of fibrils is not seen in allo- or autografts may be due to an increased synthesis of small proteoglycans, particularly decorin and hyaluronan, preventing large collagen fibril formation (Scott 1990; see below).

It should be also mentioned that the conclusions drawn in this discussion are based on the assumption that large collagen fibrils of the allograft tendons do not undergo a process of disaggregation into smaller fibrils similar to that described by glycerol treatment of mature collagen fibrils, which is reversible (Leonardi et al. 1983). This is a possibility which must be seriously considered but is very difficult to verify experimentally without rigorous immuno-electron microscopy.

The remodelling of collagen in tendon auto-, allo- and xenografts has been elegantly quantified by the now classic work of Klein et al. (1972). They noticed at 3 months that xenografts lost 99%, allografts lost 63% and autografts 54% of their original collagen, demonstrating a clear antigenic influence on collagen turnover which was, however, still substantial—even in the autografts at 3 months. The remodelling of collagen during medial ligament repair in the rabbit has also been shown to be prolonged in that the collagen concentration takes many months to approach normal levels (Klein et al. 1972) and appears to be similar in larger animals (Woo et al. 1987a).

The observations and conclusions in this auto/allograft review throw into question the current
timeframes for rehabilitation. It is generally concluded by most knee surgeons that the graft tissue will eventually mature, given enough time, and that graft tensile strength will also increase with time, especially if the athlete waits for up to 1 year for graft maturation. The observations in this study do not support this notion that graft tensile strength will gradually increase with time. However, collagen cross-link maturation could be very important in these grafts and this may be a mechanism for restoring some graft tensile strength even when the collagen fibrils remain of small diameter. Observations by Butler et al. (1987) that primate ACL autografts and ACL allografts using the patellar tendon were only about 30% the strength of normal ACL at 12 months postimplantation, strongly support the quantitative fibril observations outlined in this human auto/allograft study.

Comprehensive detailed clinical studies by Shelbourne (Shelbourne & Nitz 1990; Shelbourne & Davis 1999), and more recently by Barber-Westin et al. (1999), have demonstrated (contrary to expectations with the preceding ACL graft small fibril data) that early accelerated ACL rehabilitation programmes have not led to increased AP laxity, which is a paradox when compared to the results of Beynon et al. (1997). It may be speculated that the early aggressive ACL rehabilitation advocated by Shelbourne may limit the collagen loss and may assist in early collagen fibril orientation and deposition, and hence may help to explain the lack of AP laxity with these programmes.

Excellent comprehensive reviews of all aspects of human ACL reconstruction are recommended further reading of this complex topic (Frank & Jackson 1997; Fu et al. 1999, 2000). ACL graft selection has been recently reviewed by Bartlett et al. (2001).

Goat ACL patellar tendon autograft collagen remodelling: quantitative collagen fibril analyses over 3 years

The aim of this study (Ng et al. 1995a, 1995b, 1996a, 1996b) was to quantify in detail the collagen fibril remodelling process in adult goat patellar tendon ACL autografts over a 3-year timeframe. Eleven mature female adult goats were used in this study. The middle one-third of the right patellar tendon was harvested and used as an ACL graft. The animals were sacrificed at the following time intervals: 6 weeks \((n = 3)\), 12 weeks \((n = 2)\), 24 weeks \((n = 2)\), 52 weeks \((n = 3)\) and 3 years \((n = 1)\). The ACL grafts were then obtained and prepared for quantitative ultrastructural collagen fibril analyses.

The normal ACL \((t = 0, n = 5)\) and ACL patellar tendon grafts \((n = 11)\) were divided into thirds and 1 mm thick sections were cut from the femoral, middle and tibial thirds. This section was then cut into a strip and four sections obtained: two were deemed superficial and contained a synovial surface and two were deemed deep. The collagen fibril profiles were directly quantified from electron micrograph negatives using a specifically designed software program for automated computerized image analysis. The frequency of fibrils within 20 diameter size classes and the percentage area occupied for each diameter group of fibrils were automatically calculated as a mean (Fig. 4.31).

The results of the study were as follows:

1 Normal adult goat ACL, \(t = 0\). The distribution was clearly bimodal with a large number of small fibrils \(< 100 \text{ nm}\) in diameter and a group of larger fibrils \(> 100 \text{ nm}\) in diameter (Fig. 4.32a). A small number of large fibrils of \(> 100 \text{ nm}\) contributed about 45% of the total collagen fibril area; these large fibrils seen in the adult goat ACL are not seen in the normal adult human ACL.

2 Normal adult goat patellar tendon, \(t = 0\). The collagen fibril distribution was quite different to the ACL with a more unimodal distribution. There were less small-diameter fibrils and more larger diameter fibrils \(> 100 \text{ nm}\) than in the ACL, and these latter fibrils contributed 65% of the total collagen fibril area (Fig. 4.32b).

3 Patellar tendon ACL graft: tibial region versus femoral region (Figs 4.33 and 4.34). At 6 weeks postsurgery there was a large increase in the number of small-diameter collagen fibrils \(< 100 \text{ nm}\) which was greatest at the tibial end. This number...
Fig. 4.31 The methodology of sampling the goat ACL, as well as tissue blocking and the electron microscopic grid random sampling technique.

Fig. 4.32 Electron micrographs of: (a) normal adult goat ACL, and (b) normal adult goat patellar tendon.

Fig. 4.33 Summary diagram of goat ACL graft fibril profiles at 12 months postgrafting. The preponderance of small fibrils at the tibial end of the graft and some large fibrils remaining in the femoral end of the graft at 12 months postgrafting should be noted. AMB, antero-medial bundle; PLB, postero-lateral bundle.

of small-diameter fibrils was increased at both ends of the graft at 52 weeks. There was a loss of the large-diameter fibrils (> 100 nm) which could be seen at 6 weeks, and at 52 weeks only a few
large fibrils remained at the femoral region. This is reflected if the ratio of large > 100 nm fibrils to small < 100 nm fibrils is plotted (Fig. 4.35).

4 Patellar tendon ACL graft superficial region versus deep region. The major observation was the complete loss of the large-diameter fibrils at 1 year from the superficial regions of the graft. There was also a concomitant increase in the number of small-diameter fibrils in the deep region at 52 weeks not seen in the superficial regions. If one plots the ratio of large-diameter fibrils > 100 nm to small-diameter fibrils < 100 nm, the major fall in the ratio is seen in the superficial region reflecting the complete removal of the large-diameter fibrils in the superficial regions of the grafts at 52 weeks (Figs 4.32, 4.35 and 4.36).
The collagen fibril profile for each time group was compared to the control ACL and patellar tendon with Kolomogorov–Smirnov analyses. Differences were found between the patellar tendon collagen profile with the ACL grafts with the 12-, 24- and 52-week groups ($P = 0.005, 0.07$ and 0.1, respectively). All the grafts except the 6-week group contained mainly small fibrils ($< 100 \text{ nm}$). The large fibrils were not repopulated in the 1-year grafts, but in the one graft studied at 3 years large fibrils $> 100 \text{ nm}$ were present. A positive correlation ($r = 0.6, P = 0.07$) was found between the percentage of fibrils of $> 100 \text{ nm}$ diameter with Young’s modulus of the grafts, which was tested in a separate study.

This study has determined the anatomical regions of the ACL graft that undergo remodelling and that this process continues for up to 3 years postgrafting. This remodelling process changes the collagen fibril profile of the original patellar tendon to one containing a greater proportion of small-diameter fibrils. The remodelling process occurs from the most outer areas inwards, and is more vigorous in the tibial region of the graft. This would be consistent with synovial revascularization. The rapid depletion of the large fibrils in the grafts as early as 6 weeks is also consistent with the dramatic decrease in mechanical and material properties of such grafts (Fig. 4.35). This collagen fibril study in the goat also parallels that in the human ACL patellar tendon grafts and appears to be a useful model to follow collagen remodelling in the ACL graft.

A 3-year biomechanical and viscoelastic study of patellar tendon autografts for ACL reconstruction

In this study, 27 adult female goats were tested—four served as controls and the others received an autograft to the right knee with each left knee serving as an additional control (Ng et al. 1995a, 1995b). The animals with grafts were tested at 0 weeks ($n = 4$), 6 weeks ($n = 4$), 12 weeks ($n = 4$), 24 weeks ($n = 3$), 1 year ($n = 5$) and 3 years ($n = 3$) after surgery. The AP laxity of the knee joint, load–relaxation, and structural and mechanical properties of the graft were tested.

The AP laxity was significantly greater than that of the controls for all groups except at 3 years where the AP laxity almost approached normal. The mechanism of this return of AP laxity to near normal is not known but may be due to stiffening of the collateral ligament complexes or may be due to tibial and femoral condyle remodelling. Load–relaxation was greater than that of the control ACLs, but in the 1- and 3-year grafts load–relaxation was less than that of the patellar tendons with 5 min of sustained loading. Between 12 and 52 weeks, the stiffness and the modulus of the grafts were 44 and 49% those of the control.

Fig. 4.36 Representative electron micrographs of ACL and patellar tendon autografts: (a) at 6 weeks, (b) at 12 weeks, (c) at 24 weeks, and (d) at 52 weeks postgrafting. The progressive loss of the large-diameter fibrils at $t = 0$ (see Fig. 4.32b) and the accumulation of small-diameter fibrils with increasing graft age should be noted.
ligaments, respectively; the modulus was 37 and 46% that of the control ACLs and patellar tendons, respectively (Figs 4.37 and 4.38). The persistent inferior mechanical performance at 3 years suggests that ACL grafts in the goat may never attain normal ACL strength.

**A 3-year study of collagen type and cross-links for ACL patellar tendon autografts in a goat model**

The collagen matrix provides the tensile strength of ligaments. Previous animal studies have shown that ACL patellar tendon autografts contain mainly small-diameter collagen fibrils (< 100 nm) at 1 year postsurgery. The long-term collagen fibril profile and biochemical changes are not clear. This study examined goat ACL patellar tendon autografts up to 3 years postsurgery for collagen type and hydroxypyridinium (HP) cross-link density (Ng et al. 1995a, 1996a).

Twenty-two mature female goats received an ACL patellar tendon autograft to the right knee and were tested at 6 weeks (n = 5), 12 weeks (n = 4), 24 weeks (n = 5), 1 year (n = 5) and 3 years (n = 3). Two in each group were assigned for collagen typing and HP analyses. Two normal animals served as controls for collagen typing and HP analyses.

Type III collagen analyses with SDS gel electrophoresis showed an increase from 6 to 24 weeks and then decreased afterwards. At 3 years, the
The grafts and control ACLs had comparable mean hydroxypyridinium cross-link densities but different Young’s moduli, which implies hydroxypyridinium cross-link density is not the only determinant for material strength. Other factors such as collagen fibril size, density of packing and orientation may affect the Young’s modulus. However, the good correlation on regression analysis suggests the hydroxypyridinium cross-link density may be an important indicator for ACL material strength. A previous study in MCL scars in rabbits has also demonstrated a positive relationship between hydroxypyridinium cross-link density and the failure stress of ligament scar (Frank et al. 1994, 1995). Chan et al. (1998) have also demonstrated a high correlation with pyridinoline content and failure loads with the healing patellar tendon in the rat.

These two similar observations in different species and different tissues indicates that hydroxypyridinium cross-link density is one of the important determinates for the tensile strength of ligament repair and during ACL graft remodelling. We currently do not understand the biological and mechanical factors that control collagen cross-linking and hence clinical therapies to improve and enhance collagen cross-link density are not yet available. There have been no studies on human ACL graft tissues and the hydroxypyridinium cross-link density to date.
Enhancement of ligament/tendon repair using tissue engineering technologies

Because of the long delay in achieving adequate tensile strength in ligament and tendon repairs to withstand the imposed tensile forces of daily living, researchers and clinicians have explored means to not only hasten the repair process in its three phases (as described above), but also to improve the quality of the repair scar tissue, which is known in the adult to be of poor mechanical quality. For example the normal rabbit MCL when cut surgically heals with the usual phases and a hypervascular disorganized scar tissue is left that is then remodelled over many weeks to years. Biomechanical testing of this scar tissue even at 1 year postinjury reveals it to be permanently weaker and less stiff than the normal MCL on both a structural and material basis. If the two ends of the ligament are opposed by sutures after severance and the gap minimized, the MCL can heal structurally to about 70–80% of the strength and structural stiffness of the normal MCL; but if a gap is left the healing is about 40–50% of the normal MCL. On a material basis (i.e. per CSA) in both the opposed and gap situations, the scar tissue only reaches a maximum of about 30% of normal MCL strength after years of healing (Frank et al. 1999).

The use of growth factors to attempt to enhance the healing phases of ligament and tendon has been pursued over the last decade since the isolation of growth factors such as transforming growth factor beta (TGF-β), insulin-like growth factor I (IGF-I) and IGF-II and also basic fibroblast growth factor (bFGF). These growth factors have been used both singly and in combinations using various delivery modes. The results have been rather disappointing in that more collagen can be deposited but it is not well organized and the collagen fibrils are usually of small diameter and hence does not lead to any long-term improvement of the quality of the repair tissue.

The use of mesenchyme stem cells is a useful strategy if the matrix synthetic capacity of these undifferentiated cells can be harnessed to enhance collagen deposition and its remodelling in the early phases of the healing response. The use of these cells is still in its early phases although results are encouraging (Young et al. 1998; Awad et al. 1999).

During the healing process in ligaments, Frank et al. (1999) noted several specific problems with the rabbit MCL scar tissue, which appeared to have mechanical implications (i.e. flaws, decreased collagen cross-links and persistence of small-diameter collagen fibrils). The latter has been identified in this chapter as the probable key explanation for the poor biomechanical performance of human and goat ACL auto- and allografts. The role of the small leucine-rich proteoglycans, such as decorin, fibromodulin and lumican, in controlling collagen fibrillogenesis in vitro has been well documented. Decorin stands out as an important molecule in controlling collagen fibril assembly (Vogel & Trotter 1987; Birk et al. 1995; Weber et al. 1996; Danielson et al. 1997).

Based on these and other observations, Frank and his co-workers hypothesized that inhibition of decorin expression during the early phases of ligament healing may enhance the size of newly synthesized collagen fibrils. They used an antisense method using binding of oligodeoxynucleotides to target decorin mRNA in collaboration with Nakamura et al. (1998). Remarkably, using this technology they were able to demonstrate that most antisense-treated scars contained bundles of aligned collagen with some restoration of the normal collagen crimp pattern. Even more remarkable was the appearance of large-diameter collagen fibrils in five of the six antisense-treated scars. Collagen fibril diameter analysis of average collagen fibril sizes for the antisense-treated scars, sense control scars, injection control scars and normal MCL were: 104.7 ± 51.1, 74.8 ± 11.00, 78.2 ± 11.9 and 189.1 ± 104.0 nm, respectively (Fig. 4.40). Correlated with these collagen fibril observations was an increase in the mechanical quality of the scar tissue. In other words, antisense scars were significantly (18–22%) less susceptible to elongation (creep) during low stress creep testing than both sense control and injection control scars. They failed at
14.9 ± 6.62 MPa on average, which was significantly stronger (by 83–85%) when compared to both sense control (8.07 ± 3.45 MPa) and injection control scars (8.16 ± 3.86 MPa) (Fig. 4.41) (Frank et al. 1999).

This exciting work is the first to demonstrate that collagen fibril size can be manipulated during healing and can also improve the mechanical strength of the repair. This new knowledge may translate in the future to clinical therapies which may be able to more rapidly improve ligament and tendon tensile strength during the complex phases of healing.

Acknowledgements

The author is grateful to Dr Michael Benjamin, University of Cardiff for Fig. 4.1 and to Professor A.W. Parker for the use of Fig. 4.32 and also to Mrs Sue Simpson for preparation of the diagrams in the text. Thanks also go to orthopaedic colleagues Owen Deacon and Iain Mclean and other Australian orthopaedic surgeons for their continued interest in supplying human ACL graft biopsies and for their collaboration with the goat ACL autograft work.

References

Repair of Tendons and Ligaments


Chapter 5

Tissue Healing and Repair: Bone and Cartilage

K.-M. CHAN, HENRY C.L. HO AND CHRISTOPHER W.C. TONG

Healing and regeneration of bone

Structure of bone

Bone is a specialized form of connective tissue, which serves the following purposes.

1. It is a skeletal framework for the body.
2. It allows the different parts of the body, such as the appendages, to move in space through various degrees of freedom by means of joints controlled by the activity of their attached muscles and stabilized by ligaments.
3. It protects the vital organs such as the heart, lungs and abdominal viscera from external trauma.
4. It is a source of production for the elements of blood.
5. It plays a major part in mineral metabolism.
6. The bone marrow is an important part of the immune system.

The structure of bone, like that of other connective tissues, is composed of cellular elements and their surrounding matrix, which is unique in that it has an organic as well as an inorganic mineral component. The cells include bone-forming cells (osteoblasts), bone-resorbing cells (osteoclasts) and a stable population of maintenance cells (osteocytes). The relative activity of each of these cell types depends on the physiological or pathological state of this unique tissue. Under normal circumstances, in the absence of trauma, the osteocytes are lined up neatly in layers, like the skin of an onion, in the line of stress to form lamellar bone. In cortical bone of long bone diaphysis, these layers form whorls surrounding a central vascular channel containing arterioles, venules, capillaries, nerves and lymphatics. These whorls are known as Haversian systems (Fig. 5.1). The osteocytes form connections with each other and are bathed in a limited and complex interconnecting fluid space called the canaliculi. In the metaphyseal portion of long bone, the lamellae are also stress orientated but form interlacing pillars called trabeculae, rather than Haversian systems, to support loads. Like many tissues in the body, bone is constantly undergoing remodelling in response to the loads that it is subjected to. There is a delicate balance between bone formation by osteoblasts and bone removal (resorption) by osteoclasts. The rate of

Fig. 5.1 Haversian canals consisting of well-organized layers of osteocytes interconnected by a system of canaliculi.
remodelling is more pronounced in trabecular bone and less in cortical bone.

The cells are derived from osteoprogenitor cells, which are themselves derived from immature mesenchymal cells of the inner cambium layer of the periosteum and endosteum lining the external and internal surfaces of bone, respectively. Under special circumstances, such as fractures, the rate of differentiation from osteoprogenitor cells to osteoblasts and osteoclasts is dramatically increased in an attempt to heal the fracture.

The strength of bone is derived from its matrix, which is produced and maintained by the osteoblasts and osteocytes. The matrix consists of an organic (40% of dry weight) and an inorganic (60%) component. The organic component is composed of mainly type 1 collagen fibres (90%), which are responsible for the tensile strength of bone, proteoglycans for its compressive strength, and other matrix proteins such as osteocalcin and osteonectin, which are important for the promotion of bone mineralization. Also present in trace amounts are growth factors and cytokines such as transforming growth factor (TGF-β), insulin-like growth factor (IGF) and interleukins 1 and 6 (IL-1 and IL-6), which aid the control of bone cell differentiation, activation, growth and turnover (Bostrom et al. 2000).

The inorganic component is the mineral portion consisting of calcium hydroxyapatite and osteocalcium phosphate (brushite), which are trapped in and between collagen molecules. These are responsible for the compressive strength of bone.

**Fractures in sports**

Fractures anywhere in the body have serious implications for the athlete since both the inhibition from pain and the disruption of skeletal framework significantly compromise the function of the affected part. Fractures sustained during sport are common. Templeton et al. (2000) found that sport accounted for 22.1% of all tibial diaphyseal fractures, of which 79.5% were sustained during soccer. In addition, the 10 activities that most commonly lead to injury for all ages are cycling, Australian football, basketball, soccer, cricket, netball, rugby, roller skating/blading, skateboarding and trampolining (Finch et al. 1998). Furthermore, sports-related injuries cause significant morbidity. In a retrospective study by Shaw et al. (1997) on the epidemiology and outcome of 523 tibial diaphyseal fractures in footballers, only 73.9% of the patients had unimpaired sporting function after the injury and the average time to return to football was 7–8 months. Therefore, it is unlikely for a player to return to sport in the same season. The crux of the problem lies in the optimal time to return to full sporting activity without jeopardizing the structural integrity of the bone. Fractures may occur in any part of the bone, but specific areas such as those around the major joints will be discussed in greater detail since the interplay between early return of joint motion and perfect restoration and healing of the articular surface are of paramount importance if long-term osteoarthritis is to be avoided.

**Aim and principle of fracture treatment**

Fractures commonly encountered in athletic activities, usually arise from repeated stress and are commonly treated non-operatively. An example is the metatarsal fracture, also known as a march fracture. However, fractures arising from more major injuries can result in disruptions of larger structures such as the tibial plateau. The tibial plateau fracture is a good example of an injury that can have a major impact on an athlete’s career and can indeed equate to permanent disability. The major concern here is to achieve accurate reduction of the fracture to restore a normal smooth articular surface followed by stable fracture fixation and to attain bone healing in as short a time as possible. As soon as bony integrity is achieved, weightbearing and mobilization can begin. Range of motion exercises of the joint should begin as soon as possible to prevent joint stiffness; they also have the role of promoting nutritional delivery and healing in the articular cartilage. Early safe mobilization also promotes the return of muscle function. While these are the
basic principles behind the management of more complex periarticular fractures, the optimum rehabilitation strategy is yet to be defined.

The healing process of fractures after injury

The process of healing in bone is one of regeneration of bone, in contrast to healing in other tissues such as muscle where the defective tissue may be replaced to some degree by scar tissue. Fracture healing is a complex and well-orchestrated physiological process and much research has been invested in this field to achieve a better understanding of the natural pathways so that potential avenues of intervention to promote earlier fracture healing can be exploited. This has important therapeutic implications since the main concern for the athlete is how soon he or she can return to normal sport.

To understand fracture healing, it is important to appreciate the cellular events taking place at the fracture site (Fig. 5.2). Fracture healing can be broadly divided into primary and secondary healing. Primary healing, also known as primary cortical healing, involves a direct attempt by the cortex to re-establish mechanical continuity. In secondary fracture healing, the periosteum containing cells of osteogenic potential and the external soft tissues form a callus, which subsequently bridges the fracture gap and undergoes cartilaginous changes followed by replacement with newly formed bone by the process of endochondral ossification. In the case of primary fracture healing, osteoclasts and osteoblasts function to remove bony debris and produce immature bone, respectively. This is followed by a vascular ingrowth to form the so-called ‘cutting cone’ (Fig. 5.3), which eventually leads to the re-establishment of Haversian systems. This process of primary fracture healing only takes place when there is a minimal fracture gap and rigid fracture fixation.

On the other hand, secondary fracture healing, which is the more common mode of healing found in fractures, involves an initial stage of fracture haematoma formation and the release of important signalling molecules including IL–1 and IL–6, TGF-β and platelet-derived growth factor (PDGF). The bone itself also releases PDGF, TGF-β, IGF-I, IGF-II and bone morphogenic protein (BMP). These function to orchestrate the various complex cellular changes taking place, which include the proliferation and differentiation of mesenchymal stem cells found in the periosteum into osteoblasts. During the first 10 days, these form new bone directly opposed to the cortex, a process known as intramembranous ossification. At the same time, the bridging callus, which is contributed by the periosteum and external soft tissues, undergoes cartilaginous transformation. At 2 weeks, calcification of the cartilage begins to

---

**Fig. 5.2** Callus at the fracture site consisting of osteoblasts, osteoclasts, newly formed blood vessels and immature bone and cartilage.

**Fig. 5.3** Cutting cone: an important functional unit consisting of leading osteoclasts resorbing bone debris, thus creating a pathway for trailing osteoblasts and newly formed vessels to lay down immature bone.
take place. This tissue then becomes the target tissue for vascular ingrowth delivering osteoblasts to the area. The extent of this vascular invasion depends very much on the integrity of the surrounding blood supply to the periosteum, endosteum and soft tissue. Newly formed immature woven bone is laid down while the calcified cartilage matrix is digested away by chondroclasts. With time, under the influence of loading during rehabilitation, remodelling takes place and woven bone becomes substituted by mechanically competent and stress-orientated lamellar bone.

MECHANICAL FACTORS INFLUENCING THE HEALING PROCESS

With the above knowledge in mind, one has to have some evidence on which to devise a rehabilitation strategy. It is important to recognize those mechanical factors that can promote and retard the natural biology of fracture healing. Only then can the clinician devise a well thought-out rehabilitation plan that is based on scientific rather than anecdotal evidence.

Investigators (Cunningham et al. 1998) have studied the effects of strain rate and timing in the mechanical modulation of fracture healing. They studied the effects of applying cyclic interfragmentary micromovement at high and low strain rates at various stages of healing and discovered that a high strain rate, similar to that obtained during brisk walking, induces a greater amount of callus formation during the early phase of healing but significantly inhibits healing in the later period (6 weeks after fracture). They also predict that in the later (hard callus) phase, tissue damage may occur under abnormally high stresses and strains and recommend increasing the rigidity of fracture fixation until the fracture has healed. It has been suggested that as healing progresses, the rigidity of the fixation system should be increased in order to minimize such excessive strains (Kenwright & Gardener 1998). This can be achieved in clinical practice by the process of dynamization, whereby the fracture ends are jammed together further as the patient bears weight.

In addition, the amount of strain and hydrostatic pressure applied along the calcified surface of the fracture gap have also been found to affect the type of bone healing taking place. Using animal models, Claes et al. (1998) discovered that for strains of less than 5% and pressures of less than 0.15 MPa, predominantly intramembranous ossification was seen. Strains between 5 and 15% and hydrostatic pressures greater than 0.15 MPa stimulated endochondral ossification, and larger strains led to healing by fibrous connective tissue. The amount of strain also affected the production of TGF-β. Increased production was found for strains of up to 5% and a subsequent decrease for larger strains. The amount of strain taking place at the fracture site is clearly dependent on the type of fixation. A fracture fixed with a plate is much more likely to demonstrate less strain compared with one fixed with an Ilizarov ring fixator. Similarly, Augat et al. (1996) studied the effects of early weightbearing after fracture fixation using more flexible devices in the sheep long bone. Although there was an abundance of early soft callus formation, there was a significant delay in the healing of the osteotomy site. On the other hand, prevention of early full weight-bearing and delayed full weightbearing resulted in a higher flexural rigidity of the fracture, an increased mechanical stiffness of the callus tissue, and an enhanced bone formation at the healing front (Augat et al. 1996).

The amount of stability that a fixation device imparts to a fracture will determine the optimum time to begin loading across the fractured bone. Sarmiento and Latta (1999) showed the beneficial effects of early weightbearing on fracture healing in their studies on rats and indeed confirmed Wolff's hypothesis that mechanical loading stimulates an osteogenic response in bone. The exact mechanism by which such mechanical influences act is still not defined but is certainly a fascinating area deserving further research. While recognizing the limitations of research models, one must remain critical about such data and question their validity when applied to clinical situations. For example, much of the research has been on animal experiments and the fracture pattern and
behaviour may be quite different to that occurring in humans. The surgical method and technique of fracture fixation also influence the stability of the osteosynthesis and affects the surgeon’s preferred method of rehabilitation. Therefore, the results of research experiments can only serve as a guide to the biological events taking place in a patient and cannot be strictly extrapolated.

The current practice of fracture fixation has now moved away from absolute rigid fixation using compression plates and rigid statically locked intramedullary nails to one of stable fixation allowing some micromovement to take place to stimulate early callus formation (Fig. 5.4). Indeed, some of the inherently more stable fractures such as the simple and minimally comminuted tibial shaft fractures may be treated primarily by a functional brace, allowing early weightbearing to stimulate micromotion. This method obviously spares the patient a surgical procedure, such as intramedullary nail fixation, and its surgical risks and allows a shorter rehabilitation time (Sarmiento & Latta 1999). Furthermore, the subsequent shortening and deformity found when bracing is used in the treatment of less stable tibial fractures might be functionally insignificant (Sarmiento & Latta 1999).

**BIOLOGICAL FACTORS INFLUENCING THE HEALING PROCESS**

In addition to using less rigid forms of fracture fixation, there is a growing trend towards using minimal fixation. Apart from the mechanical factors mentioned above, it is important to bear in mind that the chosen method of fixation must respect the local blood supply to the fracture because healing will not take place in the absence of an adequate blood supply due to a paucity of cellular elements and growth factors at the fracture site. The fixation device will ultimately fail if there is non-union of the fracture. Minimal fixation preserves the possibly already traumatized soft tissue envelope to the underlying bone and simultaneously allows more micromotion. The importance of a good blood supply is exemplified by the preponderance of fracture non-union in the relatively less well vascularized distal tibial shaft fractures. The much better vascularized metaphysis, on the other hand, nearly always heals uneventfully.

Therefore, in a tibial plateau fracture, the fixation method usually now involves arthroscopically assisted reduction of the articular surface and percutaneous screw fixation supplemented by an external ring fixator. Few would advocate open reduction and plating of such injuries. Indeed, if plating is necessary, such as in long bone fractures of the forearm, minimally invasive techniques are being developed to slide the plate under the soft tissue after making a much smaller surgical wound. Plates such as the low-contact dynamic compression plate have also been designed with the aim of reducing the amount of pressure applied to the underlying periosteum and bone to minimize ischaemia.

**Management of fracture**

Although surgical stabilization and anatomical realignment remain the mainstay of treatment of most displaced fractures, it is still important to realize that the ultimate goal is to achieve solid bone union as early as possible. One must have the armamentarium to treat delayed or established non-unions and the assessment of such challenging problems is a test of one’s understanding of the biological requirements of healing. Apart from surgical means of promoting union, such as bone grafting and bone marrow injection to reactivate the healing process, non-surgical methods, if found to be equally effective, would generally be the patient’s choice.

Future research is directed at investigating such newer non-surgical methods. An example of an already widely used tool is ultrasound stimulation (US). This technique was first introduced by Duarte in 1983 and its effects have been confirmed in animal and clinical studies (Duarte 1983). However, the exact mechanism is still unknown, but a stimulation of the expression of numerous genes involved in the healing process including those encoding for aggrecan, IGF and TGF-β have been demonstrated (Hadjiargyrou et al. 1983).
Fig. 5.4  (a) External fixator in situ to hold a comminuted tibial plateau fracture reduced. A minimally invasive technique of stable fracture fixation. (b) External fixator holding tibial plateau fracture reduced. (c) Sarmiento brace is useful for the treatment of tibial shaft fractures. (d) Intramedullary nail providing rigid internal fixation for tibial shaft fracture.
An increase in the formation of soft callus formation and an earlier onset of endochondral ossification have been observed (Wang et al. 1994). In one animal study, US treatment resulted in a 1.5 times increase in the mean acceleration of fracture healing (Pilla et al. 1990). Heckman et al. (1994) reported a significant reduction in healing time in their multicentre, randomized, controlled clinical trial using low-intensity ultrasound for 20 min per day during the first week after a fracture. The use of ultrasound for the healing of fresh fractures has been approved by the American Food and Drug Administration.

Similarly, pulsed electromagnetic and electric fields have also been found to accelerate fresh fracture healing and augment healing in delayed union. It is believed that an exogenous electrical field applied at the fracture site can induce a mechanotransductive effect similar to that observed when bone is subjected to mechanical stress. Its use in clinical practice began in the 1970s. Deibert et al. (1994) demonstrated a 55–299% increase in strength of healed rabbit fibula that were subjected to daily ion resonance electromagnetic field stimulation. The exact cellular pathways responsible for this effect have not been well defined, but a stimulation of the secretion of numerous growth factors such as BMP-2, BMP-4, TGF-β and IGF-II in vitro have been demonstrated. Sedel et al. (1981) have extensively investigated the use of electromagnetic fields in the treatment of delayed union and non-union. Ryaby (1998) achieved an efficacy rate of 64–87% in the treatment of non-union and the tibia. However, large controlled, randomized, double-blind clinical trials are currently lacking.

In terms of pharmacological treatment, the potential clinical application of L-dopa, parathyroid hormone, bisphosphonates and zinc compounds, as well as the many different growth factors, in promoting fracture regeneration are under extensive investigation. To cite a few examples, a study by Holzer et al. (1999) showed that parentally administered parathyroid hormone to mid-diaphyseal fractures of rat femurs increased the callus area and strength and the bone density, thus calling for clinical trials in its use in healing fractures that are slow to heal. Similarly, using a rat model, Sakai et al. (1999) applied a protein called activin, a member of the TGF-β family, topically to fibular fractures and demonstrated a dose-related increase in the callus volume and callus weight. An increase in the callus strength was also observed.

Research is currently taking place in order to discover the factors that have important signaling properties in these complex cellular events, as these may lead to new insights into the future development of new techniques for promoting fracture healing. One such example is BMP, which appears to be a subset of the TGF superfamily, and has an important role in early intramembranous ossification and in the differentiation of mesenchymal cells into chondrocytes by yet to be defined mechanisms. BMPs were discovered by Urist in 1965 (Urist 1965). They are present in many tissues, such as the kidney, peripheral and central nervous systems and the cardiorespiratory system. Johnson and colleagues demonstrated encouraging clinical results in 30 patients for the treatment of non-union and segmental bone defects using human BMP extracted from human bone incorporated into allografts to form a composite (Johnson & Urist 2000). However, large clinical trials are still lacking and the doses of BMP in these trials are far higher than that found in bone, leading to fears of inducing malignant change.

With so much still unknown to us, formulating the perfect rehabilitation programme for a particular fracture is very difficult. It is still very much subject to the orthopaedic surgeon’s personal training and experience and the behaviour of the fracture as well as the patient’s compliance. Understanding the complex cellular and biomechanical processes of fracture regeneration helps us institute rehabilitation based on concrete scientific evidence. However, there is no cookbook approach to any single fracture. The interpatient biology is bound to be different and, therefore, the healing ability is likely to be different as well. Thus, the progress of any fracture healing is still likely to be monitored by radiological means. However, the correlation between radiological progress and actual mechanical strength of the bone may be difficult to establish. This, of
course, poses further problems when one is designing clinical studies comparing different methods of rehabilitation. Furthermore, the forces acting through any part of the body at any one time cannot be measured easily. Such uncertainties are reflected by the numerous opinions available regarding the rehabilitation of a fracture. For example, for a given lower limb fracture, one surgeon may advise non-weightbearing mobilization whilst another may advise touch-down walking. To date, such advice cannot be supported by any sound scientific evidence and to conduct well-controlled prospective studies presents a major scientific challenge.

Healing and regeneration of articular cartilage

Structure of articular cartilage

Articular cartilage plays a vital role in the function of the musculoskeletal system by allowing almost frictionless motion to occur between the articular surfaces of a diarthrodial joint. Furthermore, articular cartilage distributes the loads of articulation over a larger contact area, thereby minimizing the contact stresses, and dissipating the energy associated with the load. These properties also allow the potential for articular cartilage to remain healthy and fully functional throughout the decades of life despite the very slow turnover rate of its collagen matrix.

The composition of articular cartilage consists of cells (chondrocytes), matrix proteins including collagen and proteoglycans, non-collagenous proteins, water and electrolytes.

Chondrocytes

The basic building blocks of the articular surface are the chondrocytes, which constitute around 5–10% of the wet weight of articular cartilage. Chondrocytes originate from undifferentiated mesenchymal marrow stem cells. These cells in turn progress through the calcified cartilage zone to become chondroblasts. The chondroblasts become chondrocytes when they become isolated in lacunae, and the chondrocytes receive their nutritional support from the surrounding synovial fluid. Chondrocytes in skeletally mature articular cartilage do not divide but still remain alive via the glycolytic anaerobic metabolism pathway. As chondrocytes age, they exhibit a decrease in cellular activity, especially in the production of collagen and proteoglycan. The function of chondrocytes is to maintain the correct internal milieu of articular cartilage. They synthesize and maintain the various components of the matrix of articular cartilage. This includes the production of collagen, proteoglycans and non-collagenous proteins as well as enzymes (Mankin et al. 1994). They maintain the balance of synthesis and degradation of the protein macromolecular complex (Buckwalter & Mankin 1997).

Collagen

Collagen contributes around 10% of the wet weight of articular cartilage. The predominant collagen type in articular cartilage, accounting for 90–95%, is type II collagen, while types VI, XI, X and XI are also found to lesser extents. Collagen is spread throughout the ground substance in the extracellular matrix and is formed by three polypeptide chains that are cross-linked covalently (Eyre 1980). Collagen fibres do not have a large resistance to compression due to their slenderness ratio. However, they are very strong in tension and are the primary component of cartilage, responsible for providing its tensile properties (Roth & Mow 1980) Articular cartilage is composed of distinct layers where the collagen fibres have different patterns of orientation (Fig. 5.5). These layers include the superficial zone, middle zone, deep zone, tidemark and calcified zone. In the superficial zone, collagen fibres are arranged in a parallel fashion relative to the joint surface. In the middle zone, collagen fibres are arranged in a criss-cross oblique fashion. In the deep zone collagen fibres are arranged perpendicular to the joint surface (Redler 1974). The tidemark is the boundary between the deep zone and the zone of calcified cartilage.
Proteoglycans

Proteoglycans contribute around 10% of the wet weight of articular cartilage. Proteoglycans are composed of a protein core to which are attached many extended polysaccharide units called glycosaminoglycans (Muir 1983). The two main glycosaminoglycans found in articular cartilage are chondroitin sulphate and keratan sulphate. The vast majority of proteoglycans found in articular cartilage are known as aggregans, which bind with hyaluronan to form a large proteoglycan aggregate (Fig. 5.6). Chondrocytes produce aggrecan, link protein and hyaluronan, which are secreted into the extracellular matrix where they aggregate spontaneously (Muir 1983). The proteoglycan molecules contain repeating sulphate and carboxylate groups along their chains. These groups become negatively charged when placed in an aqueous solution. This negative charge attracts water molecules, which exerts a large swelling force.
pressure and thus a tensile stress on the surrounding collagen network (Maroudas 1976).

The proteoglycan molecules are the main components in articular cartilage, providing resistance to compression forces within the articular cartilage. The balance of expanding total swelling pressure exerted by the proteoglycans and the constraining tensile force developed within the surrounding collagen network determines the degree of hydration in cartilage. A disruption of this balance, resulting from damage to the articular surface collagen network, has been shown to cause increased tissue hydration and significantly changes the ability of the articular cartilage to bear loads (Setton et al. 1993).

**Mechanism of injury of articular cartilage**

Direct blunt trauma, indirect impact loading or torsional loading of a joint can damage the articular cartilage and the calcified cartilage–subchondral bone region without disrupting the surrounding soft tissues. Examples of direct blunt trauma include falling from a height or striking a joint with a large hard object. Examples of indirect impact and torsional loading include a blow to a bone that forms the subchondral part of a joint or severe twisting of a joint that is loaded. Unfortunately, these injuries occur frequently in people taking part in sports and are hard to diagnose. The amount of cartilage damage from a given load depends on how rapidly this load is applied. Slowly applied loads and suddenly applied loads differ considerably in their effects.

As mentioned previously, the articular cartilage extracellular matrix consists of water and a macromolecular framework formed primarily of collagens and large aggregating proteoglycans. The collagens give the tissue its form and tensile strength, and the interaction of aggregating proteoglycans with water gives the articular cartilage its stiffness to compression, resilience and probably its durability. When the cartilage surface is loaded, fluid movement occurs within and out of the cartilage matrix. This movement of fluid serves to dampen and distribute loads within the cartilage and to the subchondral bone (Mow & Rosenwasser 1988). When a load is applied slowly to the cartilage, the fluid movement within the cartilage allows the cartilage to deform and decreases the force applied to the matrix macromolecular framework. When a load is applied too rapidly, as occurs with a sudden impact or torsional joint loading of the joint surface, it may rupture the matrix macromolecular framework, damage cells and exceed the ability of articular cartilage to prevent subchondral bone damage by dampening and distributing loads.

Experimental studies have shown that blunt trauma can damage articular cartilage and the calcified cartilage–subchondral bone region, while leaving the articular surface intact (Donohue et al. 1983). Physiological levels of joint loading do not appear to cause joint injury, but impact loading above that associated with normal activities, but less than that required to produce cartilage disruption, can cause alterations of the cartilage matrix.

**Types of chondral and osteochondral injury and their potential for healing**

The response of articular cartilage to injury and the subsequent healing response incurred depends to a large extent on the type of cartilage injuries. Cartilage injuries can be divided into three main types: (i) chondral damage without visible tissue disruption; (ii) disruption of the articular cartilage alone with sparing of the underlying subchondral bone (an example of this would be the chondral flap injury); and (iii) articular cartilage injury involving the underlying subchondral bone (i.e. osteochondral fractures). The intensity and rate of muscle contraction affect the transmission of force to the articular cartilage. Age and genetic factors may influence the type of articular cartilage injury sustained by a particular individual (Buckwalter et al. 1996).

**Matrix and cell injuries**

These are injuries where there is a damage to the chondral matrix and/or cells and/or subchondral bone without visible disruption of the
articular surface. The proteoglycan content may decrease with a concomitant disruption of the collagen fibril network. If the basic matrix structure remains intact and enough viable cells remain, the cells can restore the normal tissue composition. However, if there is significant damage to the matrix or cell population, or if the tissue sustains further damage, then the lesion may progress to cartilage degeneration.

**Classification of articular cartilage injuries**

The grading system devised by Outerbridge (1961) is a simple working tool for describing chondral lesions (Fig. 5.7). In grade I, the articular surface is swollen and soft and may be blistered. Grade II is characterized by the presence of fissures and clefts measuring less than 1 cm in diameter. Grade III is characterized by the presence of deep fissures extending to the subchondral bone and measuring more than 1 cm in diameter. Loose flaps and joint debris may also be noted. In grade IV, subchondral bone is exposed.

**Treatment modalities of articular cartilage injuries**

Various methods have been used to enhance the healing and repair of articular cartilage injuries. These include: (i) arthroscopic lavage and debridement; (ii) marrow stimulation techniques such as drilling, abrasion arthroplasty and microfracture; (iii) osteochondral autografting; and (iv) autologous chondrocyte implantation. The postoperative rehabilitation programme will vary depending on the method of treatment for the articular cartilage defect.

**Arthroscopic lavage and debridement**

Arthroscopic lavage and debridement have been shown to produce measurable symptomatic improvement (Hubbard 1996, Levy et al. 1996). In a study by Levy et al. (1996), it was noted that debridement for acute small lesions (average size 42 mm²) in 15 young soccer players allowed all players to return to soccer at 10.8 weeks after surgery, with six excellent and nine good results. However, the follow-up was short: only 1 year. Hubbard (1996) noted that grade IV femoral condyle lesions randomized to arthroscopic lavage versus debridement did better when debrided. The Lysholm scale score increased 28 points at 1 year and 21 points at 5 years in over 50% of patients. The defect size was not measured and
radiographic evidence of progression or follow-up arthroscopy was not performed in this study.

**Marrow Stimulation Procedures: Microfracture and Drilling**

The underlying principles of the various marrow stimulation techniques are similar. The concept is that the subchondral bone is breached (either with a drill or a microfracture pick) resulting in an inflammatory response; the resultant fibrin clot that forms will bring in mesenchymal stems cells, cytokines and growth factors, which will lead to a repair response in the articular cartilage defect. The repair material, however, is primarily fibrocartilage. Recently, the technique of microfracture has gained considerable popularity: the chondral defect is first debridged arthroscopically to a stable cartilage rim and any loose flaps of cartilage are removed. The calcified cartilage layer is removed by means of a curette paying attention not to violate the underlying subchondral bone. Arthroscopic surgical picks are then used to make multiple holes in the exposed subchondral bone. These holes are made 2–3 mm apart and their function is to allow pluripotent cells from the underlying marrow to enter the defect and initiate a repair process. The advantages of the microfracture technique is that it is a simple procedure amenable to all surgeons who perform arthroscopic work. The cost is low as no sophisticated instruments or laboratory procedures are required. The use of a microfracture pick to breach the subchondral bone rather than using a drill should result in less thermal damage and cellular necrosis.

Steadman reported his 3–5 year results using the microfracture technique (Steadman et al. 1997). He noted pain relief in 75% of patients with 20% unchanged and 5% made worse. When it came to functionality and activities of daily living and work, 67% improved, 20% were unchanged and 13% were made worse. Sixty-five per cent were able to return to strenuous work and sports.

Recommendations for postoperative rehabilitation programmes after microfracture include protected weightbearing for 6–8 weeks and passive mobilization with a continuous passive mobilization (CPM) machine for 8 weeks. The benefits of CPM in the treatment of chondral

---

**Fig. 5.7** Outerbridge classification of articular cartilage injury. Grade I: Softening of the articular cartilage. Grade II: Superficial fibrillation of the articular cartilage. Grade III: Deep fissures extending to the subchondral bone. Grade IV: Exposure of subchondral bone.
defects using microfracture have been documented by Rodrigo and Steadman (1994). Postoperative weightbearing status depends on the location of the lesion. Patellar and trochlear groove lesions may be weightbearing and can tolerate a hinged knee brace with a 30° flexion stop. Patients come out of the brace when they are not weightbearing and go into a CPM machine from 10° to 90° for at least 8 h·day⁻¹ (mostly at night). If the chondral defect is in the medial or lateral compartment, the patient is kept strictly to touch-down weightbearing, with a similar CPM protocol as used in patellofemoral lesions. The CPM machine is set at 1 cycle per minute with the largest range of motion that the patient finds comfortable. Following the 6–8-week period of protected weightbearing, patients are instructed to begin active range of motion exercises and progress to full weightbearing. No cutting, twisting or jumping sports are allowed until at least 4 months after surgery.

Osteochondral autograft transplantation

Osteochondral autograft transplantation involves the harvesting of small osteochondral plugs from the non-weightbearing portions of the knee and these plugs are transplanted into the cartilage defect in the weightbearing part of the knee. Most investigators agree that lesions < 2 cm² are most suitable so that donor site morbidity is avoided (Bobic 1996). The advantages of osteochondral grafting include a mature cartilage–bone unit transfer with rapid bone healing and a subsequent return to high-level function. The disadvantages include the technical difficulty in restoring the biconvex geometric congruity of the femoral condyle in the sagittal and coronal planes and the potential damage imparted to the cartilaginous cap of the individual plugs, dependent on the force required for insertion. Other questions regarding this technique remain unanswered. Concerns include the possible donor site morbidity relating to the harvest of the osteochondral plugs, and the fact that the tissue formed between the osteochondral plugs is fibrocartilage rather than articular cartilage.

Gambella and Glousman (1999) reported good to excellent results in a group of 150 patients treated with osteochondral grafting who had isolated medial and lateral femoral condyle and trochlear lesions. All physical examination parameters, including range of motion, effusion, tenderness and crepitation improved over time. Outcomes were adversely affected by poor mechanical alignment and patellofemoral chondromalacia. Second-look arthroscopies and histological biopsies were also performed and showed hyaline-like cartilage surfaces.

After osteochondral autograft transplantation, the patient is encouraged to move the knee in a hinged brace through an unrestricted range of motion but is strictly non-weightbearing for the first 6 weeks after surgery. After satisfactory assessment the patient can begin to gradually weightbear as tolerated. Graft healing is assessed both clinically and by 3-monthly serial cartilage-specific magnetic resonance imaging (MRI) scans. Patients are only allowed to return to full sporting activities after MRI evidence of full healing has been obtained.

Autologous chondrocyte implantation

In larger lesions, surgical resurfacing using autograft transfer can be a significant challenge. Restoration of articular surface in lesions > 2.5–3 cm² and up to 10–15 cm² can be accomplished by autologous chondrocyte implantation (ACI) (Brittberg et al. 1994, Minas & Peterson 1999). The technique of ACI involves a two-stage operation in the treatment of articular cartilage lesions. The first operation consists of arthroscopic surgery where the cartilage lesion is clearly defined and a sample of cartilage is harvested from the non-weightbearing area of the knee (usually over the lateral femoral ridge). This cartilage tissue is then sent to the laboratory where cartilage cells are grown and multiplied. The patient then undergoes a second operation, which is an open procedure around 6 weeks after the first operation. The harvested cells are then transplanted into the cartilage defect and the defect is covered with a piece of periosteum harvested from the
proximal tibia. A hyaline cartilage repair tissue eventually fills the cartilage defect. The advantages of ACI include the potential to treat larger lesions, the use of autologous tissue and the reliability of obtaining hyaline-like tissues in the lesion. Disadvantages include the high cost of the procedure, the need for staged surgeries and an arthroscopy for reimplantation, as well as the possibility that the repair tissue is at best a rather unpredictable and inconsistent mosaic of bone, fibrous, fibrocartilaginous and hyaline tissue.

Since initiating the procedure of ACI in Sweden, Peterson and colleagues have performed over 1000 ACI procedures. Peterson et al. (2000) reported on the results of their first 101 patients who were treated with ACI. The results have been encouraging.

The postoperative rehabilitation protocol for ACI is based on an understanding of the natural history of the repair tissue generated with this technique (Gillon et al. 1998, Minas & Peterson 1999). There are three phases of healing associated with ACI. The first phase (0–6 weeks) is the proliferative phase which occurs early after the cells are implanted. The second phase (7–12 weeks) involves a matrix production phase where the tissue becomes incorporated and integrated into the host. The final phase of healing is the maturation phase (13 weeks to 3 years), which can take an extended period of time as the repair fully matures with its stiffness closely resembling the surrounding articular cartilage. During the proliferative and matrix production phase, the principles of rehabilitation are early motion, which aids cellular orientation and the prevention of adhesions, protection of the graft from mechanical overload, and strengthening exercises that allow a functional gait. CPM and touch weightbearing are used during these early phases. Progression to full weightbearing is generally started between 4 and 6 weeks postoperatively. During the maturation phase, the knee is gradually loaded with increased strengthening exercises and various impact-loading activities. Following these principles during the repair maturation continuum will provide an optimum environment for the tissue to grow and mature.

References
Finch, C., Vlauri, G. & Ozanne-Smith, J. (1998) Sport and active recreation injuries in Australia: evidence...


PART 3

PRACTICAL ISSUES
Chapter 6

Physiological and Performance Consequences of Training Cessation in Athletes: Detraining

IÑIGO MUJIKA AND SABINO PADILLA

Introduction

Sports injuries are an unwanted but often inevitable component of competitive sports. Throughout their careers, most athletes will experience sport-related injuries that will keep them away from training and competition for variable periods of time. In addition to the injury-related functional impairment that the athlete will suffer, these periods of training cessation will also have a negative impact on his or her physiological status and performance level. Indeed, the principle of training reversibility or detraining implies that whereas regular exercise training brings about or preserves specific adaptations that enhance a subject’s ability to tolerate the stress factors arising from training, and ultimately athletic performance (Gollnick et al. 1984; Houston 1986; Coyle 1988; Tidow 1995), training cessation results in a partial or complete reversal of these adaptations, which compromise athletic performance (Coyle 1988; Hawley & Burke 1998; Mujika & Padilla 2000a, 2001a). Recent reviews on detraining define this term as the partial or complete loss of training-induced anatomical, physiological and performance adaptations, as a consequence of training reduction or cessation (Mujika & Padilla 2000a, 2001a). This definition differentiates the process through which a trained individual loses some or all of his or her training-induced adaptations (e.g. reduced training, training cessation, bed rest confinement) and the lost adaptations themselves, which are the outcome of that process.

Athletes themselves, coaches, physicians and all professionals implicated in the process of getting an injured athlete back on the sports field should be aware of the physiological and performance consequences of training stoppage. Keeping these in mind during all three phases of the rehabilitation process contributes to limiting the negative functional impact of the inactivity associated with the sports injury, and therefore accelerates the recovery of preinjury fitness and performance level.

This chapter focuses on the detraining characteristics of highly trained athletes. However, when scientific data for highly trained athletes are lacking or insufficient, available results on moderately or recently trained subjects that might shed some light on different detraining-related issues will also be reported. In this respect, readers should bear in mind that detraining characteristics may not necessarily be identical in these two types of population (Lacour & Denis 1984; Hawley 1987; Moore et al. 1987; Coyle 1988; Neufer 1989; Coyle 1990; Fleck 1994; Wilber & Moffatt 1994; Mujika & Padilla 2000a, 2000b, 2001a, 2001b), and that observations made on laboratory-based training/detraining paradigms may not be directly extrapolated to athletes approaching their higher limits of adaptation.

Cardiorespiratory consequences of training cessation

Athletes in general, and endurance-trained athletes in particular, are characterized by quite
impressive functional adaptations of their cardio-vascular and respiratory systems. However, it has often been reported that when the training stimulus is lacking or insufficient to maintain training-induced adaptations, such as during phases one and two of the rehabilitation process, there is a marked negative impact on these systems (Lacour & Denis 1984; Sjøgaard 1984; Hawley 1987; Coyle 1988, 1990; Fleck 1994; Wilber & Moffatt 1994; Mujika & Padilla 2000a, 2000b, 2001a).

Maximal oxygen uptake

Maximal oxygen uptake ($\dot{V}O_{2\max}$) is a function of maximal delivery and utilization of oxygen. ‘Central’ oxygen delivery depends on maximal cardiac output and maximal arterial oxygen content, whereas ‘peripheral’ extraction of the delivered oxygen is expressed as the arterial–venous oxygen difference (Sutton 1992). All ‘central’ factors implicated in the oxygen transport system are likely to be affected by training cessation and suffer some degree of detraining.

According to data reported in the literature, training cessation induces a rapid reduction in maximal oxygen uptake ($\dot{V}O_{2\max}$) in highly trained individuals with a large aerobic power (> 62 mL·kg$^{-1}$·min$^{-1}$) and an extensive training background (Moore et al. 1987), even when training stoppage lasts less than 4 weeks. Indeed, Houston et al. (1979) showed that 15 days of inactivity (7 days of leg casting followed by 8 days without training) led to 4% reductions in $\dot{V}O_{2\max}$ in well-trained endurance runners. $\dot{V}O_{2\max}$ also decreased by 4.7% in a group of endurance-trained runners who refrained from training for 14 days (Houmard et al. 1992, 1993). Coyle et al. (1986) studied eight endurance-trained subjects who stopped training for 2 or 4 weeks, and reported 6% $\dot{V}O_{2\max}$ declines during upright exercise. Even more pronounced results were observed by Mankowitz et al. (1992), who reported a 12% decline in $\dot{V}O_{2\max}$ during 14–22 days of training stoppage in a group of trained runners, and by Martin et al. (1986), who indicated that 3–8 weeks of physical deconditioning in highly trained subjects brought about a 20% reduction of $\dot{V}O_{2\max}$. Similar results have been repeatedly reported in team sport athletes. Male basketball players not training for 4 weeks once their competitive season was over showed a 13.8% $\dot{V}O_{2\max}$ reduction (Ghosh et al. 1987), and 11 college soccer players decreased their $\dot{V}O_{2\max}$ by 6.9% during 5 weeks of training cessation (Fardy 1969). Smorawinski et al. (2001) have recently shown a 16.5 and 10.4% $\dot{V}O_{2\max}$ reduction in amateur cyclists and bodybuilders, respectively, as a result of 3 days of bed rest.

Conflicting results, however, are not lacking in the exercise science literature. Fifteen distance runners were shown to maintain $\dot{V}O_{2\max}$ after 10 days of training cessation (Cullinane et al. 1986). The same was proven true for a group of trained cyclists and runners who also stopped training for 10 days (Heath et al. 1983), a group of female college swimmers not training for a similar period of time (Claude & Sharp 1991) and a group of soccer players after 3 weeks of training stoppage (Bangsbo & Mizuno 1988). These conflicting results could be partly explained by the different levels of physical activity performed by the athletes during the period of training stoppage.

Time and initial fitness level seem to determine the $\dot{V}O_{2\max}$ loss during periods of training cessation. Indeed, seven endurance-trained subjects stopped training for 84 days, and their $\dot{V}O_{2\max}$ declined by 7 and 16% in 21 and 56 days, respectively, it then stabilized at that level, which was still 17.3% higher than that of sedentary control subjects. A correlation of 0.93 was observed between $\dot{V}O_{2\max}$ in the trained state and per cent $\dot{V}O_{2\max}$ decline with inactivity (Coyle et al. 1984). Pavlik et al. (1986a, 1986b) reported on 40 observations referring to highly trained road cyclists and endurance runners who, for various reasons, stopped training for 60 days. They observed a linear decline in the athletes’ $\dot{V}O_{2\max}$ ($r = -0.55$) throughout the initial 45 days of training stoppage, with no further change thereafter. Sinacore et al. (1993) reported on five male and one female endurance-trained subjects who stopped training
for 12 weeks. $\dot{V}O_{2\max}$ declined by 17.1%, and this decline was greatest in those subjects with the highest initial $\dot{V}O_{2\max}$ values. Smorawinski et al. (2001) also observed a relationship between the initial $\dot{V}O_{2\max}$ and the decrement in its value ($r = 0.73$) in amateur cyclists and bodybuilders confined to rest in bed for 3 days. Very long-term (2 years) training cessation has also been shown to be characterized by $\dot{V}O_{2\max}$ declines of 6.3% in college badminton players (Miyamura & Ishida 1990).

These and other similar reported data (Drinkwater & Horvath 1972; Nemeth & Lowry 1984; Coyle et al. 1985; Martin et al. 1986; Allen 1989; Giada et al. 1995; Katzel et al. 1997; Nichols et al. 2000) indicate that the $\dot{V}O_{2\max}$ of highly trained athletes decreases progressively and proportionally to the initial $\dot{V}O_{2\max}$ during the initial 8 weeks of training cessation, this decline usually ranging between 4 and 20%. Most investigations, however, indicate that $\dot{V}O_{2\max}$ ceases to decline thereafter (Fig. 6.1) and remains higher than that of untrained counterparts (Fardy 1969; Coyle et al. 1984, 1985), although one study reported that their subjects’ $\dot{V}O_{2\max}$ fell down to sedentary values (Drinkwater & Horvath 1972). This partial retention of $\dot{V}O_{2\max}$ by highly trained athletes may be due to the accumulated effect of long-term training, but it might also be due to the inherent, genetically determined physiological capabilities of these athletes (Wilber & Moffatt 1994).

**Blood volume**

The observed reduction in cardiovascular function during short periods of training cessation (see Fig. 6.2) can to a large extent be attributed to a decline in blood volume, which may be apparent within the first 2 days of inactivity (Thompson et al. 1984; Cullinan et al. 1986). Houmard et al. (1992) considered that part of the reduction in $\dot{V}O_{2\max}$ observed in a group of distance runners after 14 days of training withdrawal was due to a 5.1% reduction in estimated resting plasma volume. Identical 5% decreases in plasma volume have been observed by other authors in distance runners who ceased training for 10 days (Thompson et al. 1984; Cullinan et al. 1986). Moreover, eight endurance athletes who refrained from training for 2 or 4 weeks were reported to suffer 9 and 12% declines in blood and plasma volumes during upright exercise, respectively (Coyle et al. 1986).

The mechanisms responsible for these declines have not been clearly established. However, marked reductions in plasma protein content are likely to play an important part. In a study with previously untrained individuals as subjects, 16 young males underwent 4 weeks of endurance training and heat acclimatization followed by 4 weeks of training stoppage. The latter resulted in reduced $\dot{V}O_{2\max}$ (3.6%) and blood volume (4.7%). Red cell volume decreased by 98 mL, plasma volume by 248 mL, total plasma protein by 16 g and plasma albumin by 12 g. Moreover, plasma volume changes were directly related to plasma protein dynamics, as loss of protein appeared to be responsible for 97% of the reduction in plasma volume (Pivarnik & Senay 1986). In addition, 6 days of training cessation have been shown to fully reverse the effects of short-term training (6 days) on resting plasma volume and concentrations of aldosterone and adrenaline.

---

**Fig. 6.1** Third degree polynomial regression between the number of days of training cessation and the percentage loss in maximal oxygen uptake ($\dot{V}O_{2\max}$) in athletes, according to data reported in the literature.
(epinephrine) during submaximal intensity exercise, but not those of arginine vasopressin. According to the authors of the study, these observations implicate plasma volume as a significant variable in modifying the exercise response of fluid and electrolyte hormones (Shoemaker et al. 1998). In contrast, Raven et al. (1998) reported that a reduction of daily aerobic activity (active deconditioning) for 8 weeks resulted in a significant 7% $\dot{V}O_{2\text{max}}$ decrease in eight untrained women and 11 untrained men. None the less, total blood volume (4.0%) and plasma volume (3.1%) did not decrease significantly, and the baseline haemodynamic variables analysed by the investigators were unaltered by the deconditioning period. These data suggest that some of the changes in $\dot{V}O_{2\text{max}}$ accompanying physical deconditioning may be independent of changes in total blood volume and plasma volume (Raven et al. 1998).

Heart rate

RESTING HEART RATE

Whereas resting heart rate has been shown to increase significantly after only 3 days of bed rest in endurance athletes (Smorawinski et al. 2001), it has been reported not to change in athletes following short-term (10 days) training cessation (Cullinane et al. 1986). Resting heart rate did not change in five Olympic rowers and one canoist during 6–34 weeks of training stoppage following the 1988 Seoul Olympic Games (Maron et al. 1993), but it increased by 6.4% in 2–12 weeks of injury-induced training cessation in physical education students (Bánhegyi et al. 1999).

SUBMAXIMAL HEART RATE

As a consequence of the above-mentioned reduction in blood volume, the heart rate at submaximal exercise intensities increases by approximately 5–10%. Coyle et al. (1986), observed an 11% higher heart rate during submaximal exercise after a group of endurance athletes spent 2–4 weeks away from training. Heart rate while performing submaximal exercise (75 and 90% of $\dot{V}O_{2\text{max}}$) increased by 11 beats·min$^{-1}$ in a group of endurance runners after 14 days without training (Houmard et al. 1992). Madsen et al. (1993) also observed heart rates of 6–7 beats·min$^{-1}$ higher at submaximal exercise intensities after 4 weeks of insufficient training stimulus in nine endurance athletes.

Changes in athletes’ submaximal exercise heart rates have also been reported after longer periods of training cessation. Seven endurance-trained subjects exercised at the same absolute submaximal intensity in a trained state and during 84 days of no training. Their heart rates increased from 84 to 93% of maximal values during the initial 56 days of inactivity, but stabilized thereafter. Exercise of the same relative submaximal intensity, on the other hand, elicited almost identical percentages of maximal heart rate (Coyle et al. 1985). Eleven college soccer players showed increased heart rate at submaximal exercise intensities during 5 weeks without training (Fardy 1969). This result, along with a shortened length of the cardiac isovolumetric contraction phase at rest, led the author of the study to suggest that inactivity brought about an increase in the sympathoadrenergic tone of the players. Increased submaximal heart rates were also observed in adolescent (15–18 years) female athletes as a result of 12 (Drinkwater & Horvath 1972) and 23 (Michael et al. 1972) weeks of training cessation following the track season. In addition, the latter investigators also observed a progressive 16% increase in postexercise recovery heart rate, but stated that both submaximal and recovery heart rate increases seemed to level off after the initial 7 weeks of training stoppage. In line with the above results, six college football players showed an increased heart rate during submaximal exercise after 9 weeks of training cessation following their competitive season. Their values, nevertheless, remained lower than those of sedentary counterparts (Penny & Wells 1975). Giada et al. (1998) observed that heart rates at a power output of 100 W increased from 108 to 114 beats·min$^{-1}$ in young cyclists, and from 100 to 106 beats·min$^{-1}$ in older cyclists after 2 months of training cessation.
MAXIMAL HEART RATE

The training stoppage also influences heart rate during maximal intensity exercise, increasing it by about 5–10%. Cullinane et al. (1986) measured a 5% increase in maximal heart rate in 15 runners after 10 days of training cessation. Fourteen days of training cessation also resulted in 9 beats·min⁻¹ higher maximal heart rate in a group of endurance runners (Houmard et al. 1992). Madsen et al. (1993), on the other hand, reported unchanged maximal heart rates after 4 weeks of insufficient training stimulus in nine endurance athletes. In contrast, in an 84-day training cessation follow-up study on seven endurance-trained subjects, maximal heart rate increased by 4–5% after the initial 12–21 days without training, but it did not change thereafter (Coyle et al. 1984).

Stroke volume

The reduced maximal aerobic capacity observed in highly trained subjects after short periods of training cessation seems to be a consequence of the decline in stroke volume arising from the reduced blood volume that characterizes detraining (Coyle et al. 1984). Coyle et al. (1986) observed that stroke volume fell by 12% during upright exercise in a group of endurance athletes who stopped training for 2–4 weeks. This effect was reversed after plasma volume expansion, but this intervention did not restore $\dot{V}O_{2\text{max}}$, which remained 3% lower than in the trained state. This observation suggests that the reduction in blood volume associated with detraining limits ventricular filling during upright exercise and is largely responsible for the inactivity-induced decline in cardiovascular function. Stroke volume declined by 10% after the initial 12 of 84 days without training in seven endurance-trained subjects, and averaged 10–14% below trained levels during 12–84 days of training stoppage. These values did not differ from the controls. The authors indicated that the decline in $\dot{V}O_{2\text{max}}$ during the initial 21 days of training cessation was associated with a decreased stroke volume (Coyle et al. 1984). The same group reported that 3–8 weeks of physical deconditioning in highly trained subjects, which brought about a 20% reduction of $\dot{V}O_{2\text{max}}$ resulted in a significant 17.2% reduction in stroke volume during upright exercise (Martin et al. 1986). On the other hand, Pavlik et al. (1986a, 1986b) indicated that the resting stroke volume index (mL·beat⁻¹·m⁻²) and ejection fraction rose in a group of highly trained road cyclists and endurance runners during 60 days of training cessation. These changes were paralleled by a linear decrease in $\dot{V}O_{2\text{max}}$ until the 45th day of training withdrawal. In contrast, Smorawinski et al. (2001) recently reported on a significant decrease in resting stroke volume in a group of endurance athletes after 3 days of bed rest, but not in bodybuilders. The authors did not elucidate whether these differences depended on the subjects’ habitual physical training mode and endurance activity level or the genetic factors associated with aerobic capacity.

Cardiac output

Cardiac output decreases during exercise in highly trained athletes during periods of training cessation. This decrease takes place because the above-mentioned increase in exercise heart rate is not enough to counterbalance the decline in stroke volume. Coyle et al. (1984) reported that estimated maximal cardiac output stabilized at 8% below trained values after 21 days of training cessation in endurance athletes, not falling beyond that level between the 21st and 84th days without training. These authors also observed a progressive shift from 84 to 94% of maximal cardiac output between the trained and detrained (84 days without training) states when subjects exercised at the same absolute submaximal intensity, whereas exercise of the same relative submaximal intensity elicited almost identical percentages of cardiac output (Coyle et al. 1985). Slightly but significantly higher cardiac outputs were also elicited during submaximal supine exercise following 8 weeks of training cessation in highly trained subjects (Martin et al. 1986).
Detrained cyclists’ and runners’ cardiac index (L·m⁻²·min⁻¹) increased at rest as a result of the increased stroke volume index (Pavlik et al. 1986a, 1986b). Also, resting cardiac output of 55 physical education students not training for 2–12 weeks due to injury increased by 10.1% (Bánhegyi et al. 1999). On the other hand, submaximal and maximal cardiac output was unchanged in endurance athletes and bodybuilders as a result of 3 days of bed rest (Smorawinski et al. 2001).

Cardiac dimensions and circulation

Short-term training cessation has also been shown to result in altered cardiac dimensions in highly trained athletes (Fig. 6.2). Martin et al. (1986), for instance, reported significant reductions in left ventricular end-diastolic dimension (11.8%), left ventricular wall thickness (25.0%) and increased mean blood pressure during upright exercise in athletes who did not take part in physical training for 3 weeks. These investigators attributed these changes to a reduction in left ventricular mass (19.5%). The increased mean blood pressure during upright exercise measured by these and other authors (Coyle et al. 1986) could indeed be due to a reduced left ventricular mass, coupled with a higher total peripheral resistance, which increased by 8% after 2–4 weeks of training cessation (Coyle et al. 1986). Cullinane et al. (1986), on the other hand, did not observe any change in the cardiac dimensions and the blood pressure of 15 distance runners following 10 days of training stoppage. During longer periods of training cessation (8 weeks), the left ventricular end-diastolic dimension of highly trained athletes declined in parallel with stroke volume when performing upright exercise (Martin et al. 1986). Meanwhile, left ventricular posterior wall thickness decreased progressively by 25%, but left ventricular mass was unaltered after the decline observed following the first 3 weeks of deconditioning.

Six Olympic-level athletes (five rowers and one canoeist) were shown to decrease their maximal left ventricular wall thickness by 23% (range 15–33%) and their left ventricular mass by 22%.

![Cardiac morphology and function changes](image_url)

**Cardiac morphology**
- Left ventricular end-diastolic dimension (4–12%, 21 days)
- Left ventricular wall thickness (7–33%, 21 days)
- Left ventricular mass (8–37%, 21 days)
- Interventricular septal wall thickness (7%, 60 days)

**Cardiac function**
- Resting heart rate (0–7%, 14 days)
- Submaximal heart rate (5–10%, 10–14 days)
- Maximal heart rate (5–10%, 10–14 days)
- Recovery heart rate (7–16%, 21 days)
- Resting cardiac output (10–30%, 14–21 days)
- Submaximal cardiac output (5–12%, 12–21 days)
- Mean blood pressure (8–12%, 21 days)
- Total peripheral resistance (8%, 14–28 days)
- Exercise stroke volume (10–18%, 12 days)
- Maximal cardiac output (8–10%, 12–21 days)
- Blood and plasma volume (5–12%, 2 days)

Fig. 6.2 Morphological and physiological changes most likely to occur at the cardiac level, and minimum reported time of training cessation necessary for these changes to take place in athletes.
(range 8–37%) during 6–34 weeks of training stoppage following the 1988 Seoul Olympic Games. No changes in blood pressure or resting heart rate were observed. It was concluded that deconditioning may be associated with a considerable reduction in ventricular septal thickness in elite athletes, suggesting that this population had a physiological form of left ventricular hypertrophy induced by training (Maron et al. 1993). Studying a group of young (24 ± 6 years) and a group of older (55 ± 5 years) amateur cyclists during 2 months of training stoppage, Giada et al. (1998) reported a reduced posterior wall thickness index (6.6%) and interventricular septal wall thickness index (6.5%) in the young athletes. In the older athletes, detraining was characterized by a reduced left ventricular end-diastolic diameter index (3.4%), left ventricular end-diastolic volume index (9.9%) and left ventricular mass index (15.7%). The authors concluded that in the young athletes, thickness was greatly reduced whereas ventricular cavities remained unchanged, the final result being a slight and non-significant reduction in left ventricular mass. Conversely, in the older athletes, the significant reduction in left ventricular mass was the result of reduced cardiac cavities, whereas thickness of the free wall and septum remained unchanged. It was suggested that in the elderly subjects cardiac adaptation to aerobic training takes place mainly through a higher degree of diastolic filling of the left ventricle with a greater utilization of Starling’s mechanism, which makes up for the reduced chronotropic capacity.

Giannattasio et al. (1992) reported that former professional sprint runners and hammer throwers had left ventricular end-diastolic diameter and mass indexes similar to those of sedentary subjects after 4–5 years of training cessation. They also indicated that the impairment of the cardio-pulmonary reflex that they observed in athletes is to a great extent reversible with training cessation-induced regression of cardiac hypertrophy. Pavlik et al. (1986a, 1986b) observed that diastolic relative muscular wall thickness, interventricular septum thickness and left ventricular posterior wall thickness remained unchanged and above control values in road cyclists and endurance runners during 60 days of inactivity. The same was true for end-diastolic and end-systolic volumes. Small and opposite non-significant changes were observed in relative left ventricular end-diastolic volume (which tended to increase) and relative left ventricular end-systolic volume (which tended to decrease), resulting in a linearly increased ejection fraction until the 30th day of no training \( r = 0.40 \), which reached control level. The stroke volume index also increased until the 40th day \( r = 0.40 \). Since the resting heart rate remained low throughout the follow-up period, the resting cardiac index grew linearly until day 40 \( r = 0.47 \), reaching the control level by the third week, then exceeding it. No further change was observed between the 45th and 60th days. Left ventricular mean circumferential shortening velocity followed a similar increasing trend until day 30 \( r = 0.44 \). Mean arterial blood pressure remained practically stable throughout the study, but relative total peripheral resistance declined continuously \( r = -0.45 \). Based on their observations, these authors concluded that the cardiovascular regulation undergoes a peculiar shift during training cessation in that a persisting cardiac enlargement and bradycardia is associated with a temporarily unstable autonomous control due to relatively high levels of both sympathetic and parasympathetic activity. This imbalance often leads to a hyperkinesis-like syndrome when an athlete stops endurance training abruptly. These conclusions coincide with a clinical entity known as ‘detraining syndrome’ (also referred to as ‘relaxation syndrome’), which arises when athletes with a long endurance-training history suddenly abandon their regular physical activity (Mujika & Padilla 2000a). It is characterized by a tendency to dizziness and fainting, non-systematic precordial disturbances, sensations of cardiac arrhythmia, extrasystolia and palpitation, headaches, loss of appetite, gastric disturbances, profuse sweating, insomnia, anxiety and depression (Israel 1972; S’Jongers 1976).

Bánhegyi et al. (1999), on the other hand, studying a group of 22 female and 23 male physical
education students who were forced to stop exercising due to injury, reported slight reductions in relative left ventricular wall thickness and relative left ventricular muscle mass (2.6 and 2.7%, respectively). They also observed increased resting heart rate (6.4%) and cardiac output (10.1%). When comparing these results with their own previous observations on highly trained cyclists (Pavlik et al. 1986a, 1986b), they concluded that the earliest modifications in echocardiographic parameters are likely to appear in the autonomous regulation of the heart, first in sympathetic activity, then in parasympathetic tone. The change in morphological parameters, if any, would appear much later. This time course would be the result of the fast response of the nervous regulatory component, and it would depend on the previous conditioning level of the athlete, as athletes of excellent condition not training for several weeks were only shown to increase resting sympathetic autonomous activity, whereas less well-trained athletes’ changes in the components of cardiac function occur sooner and in a faster sequence (an elevation of resting sympathetic activity becomes manifest in an increased resting cardiac output, a drop in resting parasympathetic tone in a faster heart rate, and also some of the morphological parameters may show modifications).

Interestingly, three out of 10 highly trained masters athletes (59 ± 8 years) developed new asymptomatic ischaemic-appearing, exercise-induced ST segment depression on their exercise electrocardiogram (ECG) during 3 months of training stoppage (Katzel et al. 1997). After 5–8 years of follow-up, two of the three athletes with silent ischaemia experienced major cardiovascular events (sudden death, cardiac bypass surgery), whereas the other seven athletes did not have any cardiovascular events. These observations led the authors to suggest that silent ischaemia after a 3-month period of training stoppage in highly trained older athletes may be a predictor of future cardiac events (Begum & Katzel 2000).

Mean and systolic blood pressures have been shown to increase along with total peripheral resistance during 9–12 weeks without training in six endurance cyclists and runners (Martin et al. 1986) and in six college football players (Penny & Wells 1975). The values of the latter, however, remained lower than those of sedentary counterparts. It should also be noted that there have been reports of unchanged blood pressure in seven young female track athletes (aged 14–17 years) following 12 weeks of training cessation at the end of their competitive season (Drinkwater & Horvath 1972), and in Olympic rowers and canoeists not training for 6–34 weeks (Maron et al. 1993).

**Ventilatory function**

According to several reports in the literature, ventilatory function suffers a rapid deterioration when highly trained athletes stop exercising. Maximal exercise ventilation decreased in parallel with $\dot{V}O_{2\text{max}}$ (2–7%) in six runners after 15 days without training (Houston et al. 1979). Also, male basketball players not training for 4 weeks after their competitive season showed a deterioration of cardiorespiratory efficiency, as indicated by reduced maximal ventilation (9.3%), $O_2$ pulse (12.7%) and increased ventilatory equivalent (3.9%) (Ghosh et al. 1987). In the above-mentioned investigations, maximal ventilatory volume decreased in parallel with $\dot{V}O_{2\text{max}}$. A much higher decline in maximal ventilation (21.3%) has been observed in endurance athletes after just 3 days of bed rest (Smorawinski et al. 2001). Cullinane et al. (1986) did not observe a decreased maximal ventilatory volume in male long-distance runners after 10 days of inactivity, but they did observe a significantly lower maximal $O_2$ pulse.

Reports on the negative effects of longer periods of training cessation on ventilatory volume during maximal exercise have also been published. Indeed, maximal ventilation decreased by 10, 10.3 and 14.5% in 11 college soccer players in 5 weeks (Fardy 1969), young female track athletes in 12 weeks (Drinkwater & Horvath 1972) and five badminton players in 2 years (Miyamura & Ishida 1990), respectively. In addition, the latter athletes’ hypercapnic ventilatory respons-
Consequences of detraining was shown to increase significantly. Also, ventilatory volume shifted from 53 to 71% of the maximal after 56 days without training in endurance athletes (Coyle et al. 1985). Drinkwater and Horvath (1972) and Michael et al. (1972) also observed markedly increased sub-maximal ventilatory volume and ventilatory equivalent in female athletes during long-term training cessation.

**Metabolic consequences of training cessation**

When an athlete reduces his or her habitual level of physical training to a level which is insufficient to maintain previously acquired adaptations, metabolic detraining occurs. This is primarily characterized by marked changes in the pattern of substrate availability and utilization during exercise, which most often result in altered lactate kinetics that are detrimental to sports performance (Hawley 1987; Neufer 1989; Wilber & Moffatt 1994; Mujika & Padilla 2000a&b, 2001a).

**Substrate availability and utilization**

**Respiratory Exchange Ratio**

One of the metabolic consequences of a short period of insufficient training stimulus in athletes is an increased respiratory exchange ratio (RER) at submaximal and maximal exercise intensities (Table 6.1). Moore et al. (1987) observed in an athletic population that 3 weeks of training stoppage brought about a shift in RER from 0.89 to 0.95 at an exercise intensity of 60% of $\dot{V}O_{2\text{max}}$. In another investigation, seven endurance-trained athletes exercised at the same absolute sub-maximal intensity when they trained and during 84 days of training cessation; the RER increased from 0.93 to 1.00. Exercise of the same relative submaximal intensity elicited almost identical percentages of maximal heart rate, ventilation and cardiac output, but RER increased from 0.93 to 0.96 (Coyle et al. 1985). Also, the RERs of nine highly trained endurance athletes shifted from 0.89 to 0.91 when cycling at 75% of $\dot{V}O_{2\text{max}}$ after 4 weeks of insufficient training (Madsen et al. 1993). Maximal RER also increased from 1.03 to 1.06 after 14 days without training in a group of endurance-trained runners (Houmard et al. 1992). Finally, Drinkwater and Horvath (1972) reported higher RER values during submaximal and maximal exercise in female athletes 12 weeks after the end of the track season, and Smorawinski et al. (2001) observed higher RER values both at rest and at submaximal exercise intensities in endurance athletes confined to bed for 3 days. Taken as a whole, these results clearly indicate a shift towards an increased reliance on carbohydrate as an energy substrate for exercising muscles, concomitantly with a decreased contribution from lipid metabolism.

<table>
<thead>
<tr>
<th>Days of training cessation</th>
<th>Exercise intensity</th>
<th>Trained vs detrained RER</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>74% $\dot{V}O_{2\text{max}}$</td>
<td>0.93 vs 0.97</td>
<td>Coyle et al. 1985</td>
</tr>
<tr>
<td>14</td>
<td>100% $\dot{V}O_{2\text{max}}$</td>
<td>1.03 vs 1.06</td>
<td>Houmard et al. 1992</td>
</tr>
<tr>
<td>21</td>
<td>74% $\dot{V}O_{2\text{max}}$</td>
<td>0.93 vs 0.98</td>
<td>Coyle et al. 1985</td>
</tr>
<tr>
<td>21</td>
<td>60% $\dot{V}O_{2\text{max}}$</td>
<td>0.89 vs 0.95</td>
<td>Moore et al. 1987</td>
</tr>
<tr>
<td>28</td>
<td>75% $\dot{V}O_{2\text{max}}$</td>
<td>0.89 vs 0.91</td>
<td>Madsen et al. 1993</td>
</tr>
<tr>
<td>56</td>
<td>74% $\dot{V}O_{2\text{max}}$</td>
<td>0.93 vs 0.99</td>
<td>Coyle et al. 1985</td>
</tr>
<tr>
<td>84</td>
<td>74% $\dot{V}O_{2\text{max}}$</td>
<td>0.93 vs 1.00</td>
<td>Coyle et al. 1985</td>
</tr>
<tr>
<td>84</td>
<td>100% $\dot{V}O_{2\text{max}}$</td>
<td>0.97 vs 1.10</td>
<td>Drinkwater &amp; Horvath 1972</td>
</tr>
</tbody>
</table>
GLUCOSE AND LIPID METABOLISM

There have been several reports of a rapid decline in sensitivity for insulin-mediated whole-body glucose uptake during short-term inactivity. Burstein et al. (1985) studied insulin-stimulated glucose disposal and erythrocyte insulin receptor binding in seven endurance-trained athletes (six runners and one swimmer) after 12 h, 60 h and 7 days of training stoppage. Plasma glucose did not change throughout the study period, but plasma insulin increased significantly between the second and seventh days without training. The metabolic clearance rate of glucose, measured by the euglycaemic clamp technique, decreased from $15.6 \pm 1.8 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ at the 12 h time point after the last training bout, to $10.1 \pm 1.0 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ at 60 h, and to $8.5 \pm 0.5 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ after 7 days. The last two values were not different from those of the sedentary controls ($7.8 \pm 1.2 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$).

There was also a $21.4 \pm 1.8\%$ decrease in insulin receptor binding to isolated young erythrocytes between 12 and 60 h postexercise, apparently due to a reduced receptor number, rather than the affinity. These authors concluded that the increase in peripheral insulin action seen in trained athletes is rapidly reversed, and that modulation of in vivo insulin response by training stoppage may be at least partially mediated by changes in insulin receptor number. Five days of training stoppage were enough in endurance athletes to decrease insulin sensitivity to levels found in untrained counterparts (Mikines et al. 1989a). These authors, using the sequential hyperglycaemic clamp technique, reported that glucose-induced β-cell secretion increased slightly towards untrained levels during the same period, indicating that β-cells are subjected to an adaptation during training (Mikines et al. 1989b). Hardman et al. (1998) measured a 15.8% higher postprandial serum insulin response in endurance athletes after 6.5 days without exercise.

In a similar short-term inactivity study (6 days) in endurance runners, it was shown that glucose disposal rates after insulin infusion fell by 14.2–29.5%, despite 10.0–21.9% higher plasma insulin. Insulin clearance and muscle GLUT-4 glucose transporter protein were also reduced by 8–29 and 17.5%, respectively. These results indicated that short-term inactivity decreases insulin action in endurance runners, suggesting that a reduction in muscle GLUT-4 transporter level may play a role in the decrease in glucose disposal rates (Vukovich et al. 1996). McCoy et al. (1994) reported that 10 days of training stoppage in trained triathletes resulted in a 43.3% increase in the area under the insulin response curve during an oral glucose tolerance test, but still remained 24.3% below values of untrained control subjects. No change was observed following training cessation in the glucose response to the test, but GLUT-4 protein levels decreased by 33.2%, in parallel with a 28.6% reduction in citrate synthase activity, suggesting that glucose transport and oxidation are regulated by the muscle activity level. In keeping with the previous findings, the area under the glucose and insulin curves increased, respectively, by 65 and 73% during 7–10 days of training cessation in endurance-trained athletes. Resting metabolic rate fell by 4%, and the RER during the oral glucose tolerance test increased by the same amount. However, no change was observed in the calf and forearm blood flow. These data indicated a deterioration in glucose tolerance and energy metabolism, but these changes did not seem to be mediated by limb blood flow (Arciero et al. 1998).

Heath et al. (1983) also studied the effects of 10 days of training cessation on glucose tolerance and insulin sensitivity in a group of trained cyclists and runners. Glucose tolerance was significantly reduced, as shown by a 10–25% increase in blood glucose concentration and a 55–120% increase in plasma insulin concentration in response to an oral glucose tolerance test. A single bout of exercise, however, returned postglucose tolerance test glucose and insulin levels to almost initial trained values, giving support to the authors’ hypothesis that residual effects of the last bouts of exercise play an important role in the trained individuals’ improved glucose tolerance and blunted insulin response to a glucose load. Following 14 days of training cessation, an increased area under the glucose curve...
(14.8%) was observed in 12 endurance runners. The insulin curve also increased in endurance (30.3%) and strength (23.3%) in athletes during an oral glucose tolerance test. The insulin sensitivity index decreased by 23.7% in the former, and 16.0% in the latter. The authors, however, observed no change in GLUT-4 content, either in endurance runners or in weightlifters. This lack of change was evident despite a 25.3% decrease in citrate synthase activity in the endurance runners. It was concluded that the decrement in insulin sensitivity with training cessation was not associated with a decrease in GLUT-4 protein content, and that muscle oxidative capacity does not necessarily change in tandem with GLUT-4 protein content (Houmard et al. 1993).

On the other hand, 2 weeks of training cessation from endurance running has been shown to yield a condition that favours the storage of adipose tissue, as shown by a marked increase (86%) in adipose tissue lipoprotein lipase activity, coupled with a marked decrease (45–74%) in muscle lipoprotein lipase activity (Simsolo et al. 1993). Moreover, endurance-trained subjects in the fasted state have been reported to increase their concentrations of triacylglycerol (47%), very low-density lipoprotein cholesterol (28.2%) and the ratio of total cholesterol : high-density lipoprotein cholesterol (7.5%) during 6.5 days without exercise. Concentrations of non-esterified fatty acids, on the other hand, decreased during the same period by 43.5%. Due to a higher rate of triacylglycerol removal, probably related to their high lipoprotein lipase activity because of their well-vascularized skeletal muscle mass, endurance-trained subjects usually exhibit low levels of postprandial lipaemia. However, 6.5 days without training increased postprandial lipaemia by 42.2%, as shown by the increased area under the plasma triacylglycerol versus time curve, indicating that frequent exercise is necessary to maintain a low level of postprandial lipaemia in endurance-trained subjects (Hardman et al. 1998). Ten days of training cessation also resulted in a 15% reduction in high-density lipoprotein cholesterol and a 10% increase in low-density lipoprotein cholesterol in endurance-trained athletes (Thompson et al. 1984). Also, marathon runners have been shown to increase their plasma levels of lipoprotein(a) and apolipoprotein A1 by 24.2 and 16.0%, respectively, during 1 month of suspended physical activity of all types. Moreover, high-density lipoprotein cholesterol decreased by 14.6% as compared with post-marathon race values (Bonetti et al. 1995). Giada et al. (1995) studied the effects of 2 months of training stoppage in a group of young (24 ± 6 years) and a group of older (55 ± 5 years) amateur cyclists. Compared with the values recorded at the peak of their seasonal preparation, body fat increased by 4 and 2% in young and older athletes, respectively. Triglycerides (14.5 and 22.2%) and the low-density lipoprotein cholesterol : high-density lipoprotein cholesterol ratio (21.4 and 26.3%) increased significantly in both groups, whereas apolipoprotein A1 (8.4 and 8.6%), high-density lipoprotein cholesterol (10.2 and 8.7%), high-density lipoprotein cholesterol (10.5 and 12%) and high-density lipoprotein cholesterol (7.7 and 4.5%) decreased significantly. In addition, fibrinogen and very low-density lipoprotein cholesterol also increased in the group of young cyclists (5.2 and 34.0%, respectively). The authors of this study concluded that a 2-month interruption in aerobic training changed antiatherogenic lipoprotein profile and body composition unfavourably in both groups, and, therefore, that age did not seem to have a significant influence on the plasma lipid response to training stoppage.

Mankowitz et al. (1992) observed, in a group of trained runners, that 14–22 days of training stoppage resulted in a 7.7% reduction in fasting high-density lipoprotein cholesterol concentration, and a 21% reduction in postheparin lipoprotein lipase activity. No significant changes were observed in body fat, estimated plasma volume, lipoprotein(a), total cholesterol, triglycerides and low-density lipoprotein cholesterol. On the other hand, the mean areas under the concentration versus time curves for chylomicron-retinyl esters increased by 41% and for chylomicron remnant-retinyl esters increased by 37% after training cessation. It was concluded
that detraining was characterized by reduced fasting concentrations of high-density lipoprotein cholesterol and a decreased metabolism of chylomicrons—changes that are associated with an increased incidence of atherosclerosis.

During 2 months of training stoppage, six highly trained female swimmers were shown to have increased their body mass by 4.8 kg, of which 4.3 kg was fat mass. This represented an increase in body fat of about 4%. Interestingly, no change was observed in lipid and energy intake during this period, and the energy equivalent of the fat gain (170 MJ) closely matched the habitual energy cost of daily training, suggesting that the fat gain was related to a failure to reduce energy and fat intake to a level commensurate with the new energy and lipid expenditure, perhaps in order to promote the restoration of lipid balance (Alméras et al. 1997).

RESTING METABOLIC RATE

The influence of training cessation on the resting metabolic rate of athletes has not been clearly established. In one report, resting metabolic rate was shown to fall by 4% in a group of endurance-trained athletes during 7–10 days without training, indicating a reduced energy metabolism (Arciero et al. 1998). However, other authors reported that the resting metabolic rate was not affected by 3 weeks of training stoppage in trained males, which led the authors to suggest that exercise training does not potentiate resting metabolic rate (LaForgia et al. 1999).

Blood lactate kinetics

It has been clearly established that highly trained athletes respond to submaximal exercise of the same absolute intensity with higher blood lactate concentrations after only a few days of training cessation. In a recent investigation, the blood lactate threshold of endurance athletes was negatively affected after only 3 days of bed rest, falling from 71 to 60% of \( \dot{V}O_{2\text{max}} \) (Smorawinski et al. 2001). Competitive swimmers’ skeletal muscle metabolic characteristics have been reported to suffer dramatic changes affecting blood lactate kinetics in as little as 1–4 weeks of training cessation. Indeed, muscle respiratory capacity decreased by 50% after 1 week of inactivity in a group of eight swimmers. When subjects performed a standardized 183 m submaximal swim, postswim blood lactate was 2.3 times higher, pH significantly lower (7.183 vs 7.259), bicarbonate concentration 22.7% lower, and base deficit twice as high (Costill et al. 1985). A group of 24 college swimmers not training for 4 weeks after 5 months of competitive training showed a 5.5 mmol·L\(^{-1}\) increase in blood lactate concentration following a standardized 183 m submaximal swim, which was also indicative of a reduction in the muscle oxidative capacity, and/or a change in the swimmers’ mechanical efficiency (Neufer et al. 1987). Similar results were observed by Claude and Sharp (1991) in seven female swimmers who refrained from training for 10 days.

Endurance runners and cyclists performing an exercise task of the same relative submaximal intensity before and after 84 days of training stoppage showed a shift in blood lactate concentration from 1.9 to 3.2 mmol·L\(^{-1}\), along with a marked decline in the muscle respiratory capacity (Coyle et al. 1985). An increased blood lactate concentration at submaximal exercise intensity has also been reported in six college American football players after 9 weeks of postseason break (Penny & Wells 1975). Moreover, the lactate threshold has been shown to decline with 84 days of inactivity from 79.3 to 74.7% of \( \dot{V}O_{2\text{max}} \), but to remain above sedentary control values of 62.2% (Coyle et al. 1985). Pavlik et al. (1986b) reported on 40 observations on highly trained road cyclists and endurance runners who, for various reasons, stopped training for 60 days. They observed a linear decline in the athletes’ lactate threshold \((r = -0.58)\) throughout the initial 45 days of training stoppage, with no further change thereafter. Nichols et al. (2000) reported on a female masters athlete (49.5 years) who due to injury was forced to refrain from training for a period of 32 days. Compared to values assessed 2 days preinjury, when the subject was at the peak of her competitive season, power output
Muscle glycogen

Muscle glycogen concentration declines rapidly with training cessation in highly trained athletes. This decline is associated with a rapid reduction in glucose to glycogen conversion and glycogen synthase activity. In eight competitive swimmers, the deltoid muscle glycogen concentration declined by 20% in the first week of training cessation after the competitive season, and by 8–10% per week without training thereafter (Costill et al. 1985). Pre-exercise muscle glycogen has also been shown to be reduced by 20% in a group of triathletes, cyclists and runners undergoing 4 weeks of insufficient training (Madsen et al. 1993). Moreover, short-term (5 days) training stoppage has been shown to be enough in seven endurance-trained athletes to decrease glucose to glycogen conversion and glycogen synthase activity towards sedentary values (Mikines et al. 1989a).

A summary of the main metabolic consequences of training cessation in athletes can be found in Table 6.2.

Muscular consequences of training cessation

One of the most important characteristics of skeletal muscle is its dynamic nature. Skeletal muscle tissue has an extraordinary plasticity that enables it to adapt to variable states of functional demands, neuromuscular activity and hormonal signals, by reversibly changing its functional characteristics and structural composition (Saltin & Gollnick 1983; Hoppeler 1986; Kannus et al.

Table 6.2 Main metabolic changes most likely to occur, and minimum reported time of training cessation necessary for these changes to take place in athletes.

<table>
<thead>
<tr>
<th>Metabolic change</th>
<th>Percentage change</th>
<th>Minimum time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑ Maximal respiratory exchange ratio</td>
<td>3–13</td>
<td>14</td>
</tr>
<tr>
<td>↑ Submaximal respiratory exchange ratio</td>
<td>2–8</td>
<td>12</td>
</tr>
<tr>
<td>↓ Resting metabolic rate</td>
<td>0–4</td>
<td>7–10</td>
</tr>
<tr>
<td>↓ Insulin-mediated glucose uptake</td>
<td>14–46</td>
<td>2.5</td>
</tr>
<tr>
<td>↑ Postprandial insulin response</td>
<td>10–22</td>
<td>6.5</td>
</tr>
<tr>
<td>↓ Insulin clearance</td>
<td>8–29</td>
<td>6</td>
</tr>
<tr>
<td>↓ Muscle GLUT-4 protein content</td>
<td>0–33</td>
<td>6</td>
</tr>
<tr>
<td>↓ Muscle lipoprotein lipase activity</td>
<td>21–75</td>
<td>14</td>
</tr>
<tr>
<td>↑ Adipose tissue lipoprotein lipase activity</td>
<td>86</td>
<td>14</td>
</tr>
<tr>
<td>↑ Body fat</td>
<td>0–4 percentage points</td>
<td>60</td>
</tr>
<tr>
<td>↑ Postprandial lipaemia</td>
<td>42</td>
<td>6.5</td>
</tr>
<tr>
<td>↑ Triacylglycerol</td>
<td>0–47</td>
<td>6.5</td>
</tr>
<tr>
<td>↓ Non-esterified fatty acids</td>
<td>44</td>
<td>6.5</td>
</tr>
<tr>
<td>↓ High-density lipoprotein cholesterol</td>
<td>8–15</td>
<td>10</td>
</tr>
<tr>
<td>↑ Low-density lipoprotein cholesterol</td>
<td>0–10</td>
<td>10</td>
</tr>
<tr>
<td>↑ Very low-density lipoprotein cholesterol</td>
<td>28–34</td>
<td>6.5</td>
</tr>
<tr>
<td>↑ Lipoprotein(a)</td>
<td>0–24</td>
<td>30</td>
</tr>
<tr>
<td>↓ Chylomicron metabolism</td>
<td>37–41</td>
<td>14–22</td>
</tr>
<tr>
<td>↑ Submaximal blood lactate</td>
<td>150–231</td>
<td>7</td>
</tr>
<tr>
<td>↓ Lactate threshold</td>
<td>4–17</td>
<td>21</td>
</tr>
<tr>
<td>↓ Blood bicarbonate</td>
<td>8–24</td>
<td>7</td>
</tr>
<tr>
<td>↓ Muscle glycogen</td>
<td>20–39</td>
<td>7</td>
</tr>
</tbody>
</table>

at the lactate threshold decreased by 16.7%, and power output at a blood lactate concentration of 4 mmol·L⁻¹ by 18.9%. The athlete needed approximately 11 weeks to return to her pre-injury level of fitness. A meta-analysis study by Londeree (1997) has confirmed that there is a decline in the lactate threshold with training cessation.
Training-induced skeletal muscle adaptations are such that the trained muscle increases its tolerance to exercise (Houston 1986). On the other hand, as shown in Fig. 6.3, skeletal muscle tissue also readjusts to the reduced physiological stressors during periods of injury, when there is reduced training or complete training cessation (Sjøgaard 1984; Houston 1986; Hawley 1987; Costill 1988; Coyle 1988, 1990; Fleck 1994; Wilber & Moffatt 1994; Tidow 1995; Mujika & Padilla 2001b).

Muscle structure

Muscle capillarization

Given the contradictory results that have been reported in the literature, it is fair to state that the effects of training stoppage on capillary density in athletes have not been clearly established. Houston et al. (1979) reported a 6.3% reduction in capillary density after only 15 days of training cessation in well-trained endurance runners. Similarly, in semiprofessional soccer players, the number of capillaries around ST fibres decreased significantly from 6.0 to 5.8 after 3 weeks of training cessation (Bangsbo & Mizuno 1988). In contrast with these results, Coyle et al. (1984) reported that seven endurance-trained subjects' trained-state muscle capillarization was unchanged after as long as 84 days without training, and that capillarization in this population was about 50% higher than in sedentary controls. These authors discussed that the retention of the increased capillary density contributed to the observed partial maintenance of the ability to attain a high percentage of $\dot{V}O_{2max}$ without large increases in blood lactate concentration.

Muscle fibre distribution

The consequences of training cessation on muscle fibre distribution seem to depend on the duration of the inactive period. Six highly trained distance runners wore a walking plaster cast for 7 days, then refrained from training for 8 additional days, but this was not enough to induce any change in their muscle fibre distribution (Houston et al. 1979). The same was true in the case of four soccer players not training for 3 weeks (Bangsbo & Mizuno 1988), and 12 strength-trained athletes not training for 14 days (Hortobágyi et al. 1993). Longer periods of training cessation, on the other hand, have been reported to induce significant changes in the fibre distribution of athletes participating in various sports. Coyle et al. (1985) observed a large progressive shift from FTa to FTb fibres in endurance runners and cyclists, the latter increasing from 5% in the trained state to 19%...
after 56 days of training cessation. Sinacore et al. (1993) studied muscle fibre distribution in five male and one female endurance-trained subjects who stopped training for 12 weeks. They observed a 26% increase in the proportion of FTb fibres, and a concomitant 40% decrease in the proportion of FTa fibres. Larsson and Ansved (1985) reported that the proportion of ST fibres in four elite oarsmen decreased by 14–16% during the 4-year period following their retirement from competition. An opposite tendency towards a higher oxidative fibre population has been observed in strength athletes. Indeed, a case study on one elite power lifter indicated that the oxidative muscle fibre population was 1.4 times greater after 7 months without training (Staron et al. 1981), and Häkkinen and Alén (1986) reported a reduction in percentage FT muscle fibres (from 66 to 60%) in an elite bodybuilder who underwent training cessation for 13.5 months. Nevertheless, 14–15-year-old soccer players showed unchanged muscle fibre distribution after 4–8 weeks of training cessation (Amigó et al. 1998), and female dancers have also been shown not to change their fibre type distribution after long-term (32 weeks) training cessation, suggesting, beyond the possible limitations inherent to the muscle biopsy technique, that their high percentage of ST fibres may have been the result of natural selection rather than a training-induced adaptation (Dahlström et al. 1987).

**MUSCLE FIBRE SIZE**

Mean fibre cross-sectional area (CSA) has been shown to change during short-term training stoppage. Bangsbo and Mizuno (1988) studied samples of the gastrocnemius muscle of male soccer players undergoing 3 weeks of training cessation and reported a 7% decline in the mean fibre CSA. This change was primarily due to a 12.4% decline in FTa fibre area, from 6022 to 5278 μm². Similar results have been observed in 12 weightlifters, whose FT fibre CSA declined by 6.4% in 14 days of training stoppage. It is worth mentioning that plasma concentrations of growth hormone (58.3%), testosterone (19.2%) and the testosterone : cortisol ratio (67.6%) increased during the study period, whereas cortisol and creatine kinase enzyme levels decreased, respectively, by 21.5 and 82.3%. The investigators concluded, on the one hand, that short-term training stoppage in strength athletes specifically affected FT fibre size, and on the other hand, that changes in the hormonal milieu accompanying inactivity were propitious for an enhanced anabolic process, but the absence of the overload training stimulus prevented the materialization of such changes at the tissue level (Hortobágyi et al. 1993; Houmard et al. 1993). In contrast, 14 days of training stoppage did not result in a changed muscle fibre CSA in a group of endurance runners (Houmard et al. 1992); it even increased slightly (from 4.05 to 4.52·10⁴μm² in ST fibres, and from 4.20 to 5.22·10⁴μm² in FTa fibres) in a similar group of runners (Houston et al. 1979). This change can be considered a reversal of the training-induced reduction in oxidative fibre size which facilitates diffusion by reducing its distance.

Longer periods of training stoppage also brought about declines in FT and ST fibre CSA, the FT : ST area ratio and muscle mass in athletes. It has been shown that rugby league players’ CSA of FT fibres decreased to a greater extent than ST fibres, the former being 23% larger at the end of the season, but only 9% larger after 6 weeks without training. Further, an atrophy of the muscle bulk was suggested by the author in view of the fact that body mass decreased from 79.8 to 76.0 kg, but body fat content remained relatively constant during the inactive period (Allen 1989). A 49.5-year-old female masters athlete’s fat-free mass decreased by 1.8 kg and body fat increased by 2.1% during 32 days of training stoppage due to injury (Nichols et al. 2000). In line with these reports, LaForgia et al. (1999) observed a small (0.7 kg) decrease in fat-free mass during a 3-week period without training in trained males. After 7 months of training cessation, an average atrophy of 37.1% was observed in all fibre types of a power lifter, along with a large fat–weight loss (Staron et al. 1981). Also, an elite bodybuilder’s fat-free mass,
thigh and arm girth, and average fibre area decreased by 9.3, 0.5, 11.7 and 8.3%, respectively, after 13.5 months without training. In addition, the FT:ST fibre area ratio decreased from 1.32 to 1.04 (Häkkinen & Alén 1986). Häkkinen et al. (1981) also observed a reduction in the FT:ST muscle fibre area ratio from 1.11 to 1.04, and a reduced muscle mass following 8 weeks of training stoppage in strength-trained athletes, as well as decreased FT and ST fibre areas after 12 weeks without training (Häkkinen et al. 1985). Larsson and Ansved (1985) observed a 10% decrease in the relative area of ST fibres in oarsmen after long-term cessation of their athletic activity, and Amigó et al. (1998) reported a reduction in the diameter of ST and FT muscle fibres in adolescent soccer players 4–8 weeks after the competitive season. Dahlström et al. (1987), on the other hand, measured a large increase in fibre areas after 32 weeks of training cessation in female dancers, suggesting again that smaller fibres were an endurance training-induced adaptation to decrease the oxygen diffusion distance.

Muscle function

**Arteriovenous oxygen difference**

As stated above, in addition to the ‘central’ factors already discussed, \( \dot{V}O_{2\text{max}} \) is also a function of the maximal ‘peripheral’ extraction of the delivered oxygen, which is traditionally expressed as the arteriovenous oxygen difference (Sutton 1992). The authors are aware of only one study reporting data on the arteriovenous oxygen difference during training cessation in highly trained individuals (Coyle et al. 1984). According to that study, 21 days without training did not bring about any change in the maximal arteriovenous oxygen difference of seven endurance-trained athletes (15.1, 15.1 and 15.4 mL·100 mL\(^{-1}\) at days 0, 12 and 21 of training cessation, respectively). After 56 and 84 days without training, on the other hand, values fell to 14.5 (4% reduction) and 14.1 mL·100 mL\(^{-1}\) (7% reduction), respectively. Based on these results, the authors suggested that whereas the initial \( \dot{V}O_{2\text{max}} \) loss observed in athletes during training cessation is due to a decreased stroke volume, the decreased arterial–mixed venous oxygen difference would be responsible for the reduction in \( \dot{V}O_{2\text{max}} \) observed between the third and 12th weeks of training cessation (Coyle et al. 1984).

**Myoglobin concentration**

In a 12-week training cessation study, Coyle et al. (1984) reported that myoglobin concentration in the gastrocnemius muscle of seven endurance-trained runners and cyclists (43.3 mg·g protein\(^{-1}\)) did not change significantly after 3 (41.0 mg·g protein\(^{-1}\)) and 12 weeks (40.7 mg·g protein\(^{-1}\)) of training cessation. In addition, these muscle myoglobin concentrations were not different from those of eight sedentary controls (38.5 mg·g protein\(^{-1}\)). In a case study, Nemeth and Lowry (1984) reported on the myoglobin content in ST and FT fibres of the vastus lateralis muscle of a 47-year-old cyclist, at peak training and after 6 and 84 days of training cessation. The myoglobin levels remained unchanged despite a 16.7% decline in \( \dot{V}O_{2\text{max}} \) and a 15–35% decrease in the activities of oxidative enzymes during the period of training restriction. It was concluded that a change in training condition that leads to changes in the oxidative enzymes of human muscle fibres does not affect myoglobin levels of the same cells.

**Enzymatic activities**

Skeletal muscle oxidative capacity decreases markedly as a consequence of training cessation, as reflected by large reductions in mitochondrial enzyme activities. Indeed, Coyle et al. (1984, 1985) observed that seven endurance-trained subjects’ citrate synthase activity declined from 10.0 to 7.7 mol·kg protein\(^{-1}\)·h\(^{-1}\) during the initial 3 weeks of training cessation, then continued declining to 6.0 mol·kg protein\(^{-1}\)·h\(^{-1}\) by the 56th day, and stabilized thereafter. Succinate dehydrogenase, \( \beta \)-hydroxyacyl-coenzyme A (CoA) dehydrogenase and malate dehydrogenase declined roughly in parallel with citrate synthase (i.e. by about
20% in 3 weeks and 40% in 56 days of training stoppage), also stabilizing thereafter at that level, which was still 50% higher than that of sedentary counterparts. Very similar observations were made by Chi et al. (1983) after 42–84 days of training stoppage in endurance athletes—where citrate synthase, succinate dehydrogenase, β-hydroxyacyl-CoA dehydrogenase, malate dehydrogenase and β-hydroxybutyrate dehydrogenase declined by an average of 36% during that period. As in the previous study, detrained-state oxidative enzyme values were also 40% higher than controls. Interestingly, these authors also showed that mitochondrial enzyme levels decreased almost to untrained levels in ST fibres, but remained 50–80% higher in FT fibres. During 7 weeks of training followed by 3 weeks of training cessation, Moore et al. (1987) observed a 45% decline in citrate synthase activity in previously trained athletes, even though pretraining citrate synthase activity did not change in response to the initial 7 weeks of training. Also, Houmard et al. (1992, 1993) reported a 25.3% decline in citrate synthase activity in a group of 12 distance runners who stopped training for 14 days. Similar results (27% decline in citrate synthase and β-hydroxyacyl-CoA dehydrogenase) have been reported in soccer players not training for 3 weeks (Bangsbo & Mizuno 1988), triathletes (28.6% decline in citrate synthase activity) not training for 10 days (McCoy et al. 1994) and adolescent soccer players (37.5% decline in citrate synthase activity) not training for 4–8 weeks (Amigó et al. 1998). Madsen et al. (1993) reported a 12% lower β-hydroxyacyl-CoA dehydrogenase activity after 4 weeks of insufficient training in nine endurance athletes, and Houston et al. (1979) measured a 24% lower succinate dehydrogenase in six distance runners who did not train for 15 days. A similar 25% lower succinate dehydrogenase has been observed in six rugby league players’ lateral head of the gastrocnemius muscle 6 weeks after the end of their competitive season (Allen 1989). Nemeth and Lowry (1984) reported that the activities of β-hydroxyacyl-CoA dehydrogenase, citrate synthase and malate dehydrogenase fell by approximately 15–35% during 84 days of training cessation in the vastus lateralis muscle of a masters cyclist. Moreover, 16 male and female runners’ skeletal muscle lipoprotein lipase activity was reduced by 45–75% during 2 weeks of training cessation, whereas this enzyme’s activity increased by 86% at the adipose tissue level. The adipose tissue lipoprotein lipase: muscle lipoprotein lipase ratio increased from 0.51 to 4.45, which was indicative of a tendency for the storage of circulating lipids in the adipose tissue (Simso et al. 1993). Houston (1986) suggested that the observed declines in mitochondrial enzymatic activities are primarily regulated by altered protein synthesis rates. Moreover, these declines appear to be associated with the concomitant long-term reductions in $\dot{V}O_{2\text{max}}$ and arteriovenous oxygen difference (Coyle et al. 1984; Allen 1989).

Small non-systematic changes in glycolytic enzyme activities have also been reported during periods of training cessation in highly trained athletic populations. For instance, hexokinase decreased significantly by about 17%, phosphorylase did not change, phosphofructokinase increased non-significantly (about 16%) and lactate dehydrogenase increased significantly by about 20% in endurance athletes not training for 84 days (Coyle et al. 1985). Chi et al. (1983) also observed an identical 17% decline in hexokinase, whereas phosphorylase, phosphofructokinase and lactate dehydrogenase increased by 3.6–21.1% in 42–84 days without training. Houston et al. (1979), on the other hand, reported a 13% lower mean lactate dehydrogenase activity after 15 days of training cessation. Competitive swimmers not training for 4 weeks showed non-significant declines in phosphorylase and phosphofructokinase activities in their posterior deltoid muscle (Costill et al. 1985). Phosphofructokinase has also been shown to decline by 16% in rugby players not training for 6 weeks (Allen 1989), and by 54.5% in adolescent soccer players 4–8 weeks after their competitive season (Amigó et al. 1998). In addition, Mikines et al. (1989a) reported that glycogen synthase activity dropped by 42% after only 5 days without training in seven endurance-trained subjects.
MITOCHONDRIAL ATP PRODUCTION

No studies reporting on the consequences of training cessation on highly trained athletes’ mitochondrial ATP production rates seem to be available in the exercise science literature. However, using recently trained individuals as subjects, Wibom et al. (1992) reported a 12–28% decrease in mitochondrial ATP production rate during 3 weeks of training cessation consecutive to 6 weeks of endurance training. This decrease was attributed to a 4–17% reduction in individual mitochondrial enzyme activities. Because highly trained athletes’ mitochondrial enzyme activities are markedly affected by training stoppage (see above), it can be inferred that a marked reduction in mitochondrial ATP production would also take place in athletes undergoing a training cessation period. It is worth noticing, however, that the mitochondrial ATP production rate remained 37–70% above pretraining levels in the above-mentioned study on previously sedentary subjects (Wibom et al. 1992).

Training cessation and endurance performance

It has been repeatedly shown that the endurance performance of highly trained athletes suffers a rapid deterioration when the training stimulus disappears or is insufficient to maintain training-induced adaptations (Table 6.3). Smorawinski et al. (2001) have recently reported that peak power output during a graded, incremental cycle test to exhaustion declined by 14.3% in a group of endurance athletes after only 3 days of bed rest, and by 10% in strength-trained athletes. Female competitive swimmers were 2.6% slower in a 366 m swim after only 10 days without training (Claude & Sharp 1991). Exercise time to exhaustion has also been reported to be reduced during training cessation—by 9.2% (Houmard et al. 1992, 1993) and 25% (Houston et al. 1979) in 2 weeks, and by 7.6% (Coyle et al. 1986) and 21% (Madsen et al. 1993) in 4 weeks. The latter investigators suggested that altered substrate utilization and/or altered Mg$^{2+}$ transport from the extracellular to the intracellular area, which could inhibit Ca$^{2+}$ release from the sarcoplasmic reticulum, could have contributed to the reduced endurance performance after training stoppage. Interestingly, Houmard et al. (1992) suggested that the short-term training cessation-induced performance impairment observed in their 12 distance runners was primarily due to the loss in cardiorespiratory fitness they had suffered, rather than to altered mechanical efficiency, because they did not observe a decline in running economy at submaximal exercise intensities (75 and 90% of $V\text{O}_{2\text{max}}$).

Swimming performance (100 and 200 m) has also been shown to drop by 3–13% in national and international level swimmers during the inactive period in between two training seasons (Mujika et al. 1995). Also, exercise time to exhaustion has been reported to decline by 23.8% in 5 weeks

<table>
<thead>
<tr>
<th>Days of training cessation</th>
<th>Performance measure</th>
<th>Percentage decrease</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>366 m swim</td>
<td>2.6</td>
<td>Claude &amp; Sharp 1991</td>
</tr>
<tr>
<td>14</td>
<td>Exercise time to exhaustion</td>
<td>9.2</td>
<td>Houmard et al. 1992, 1993</td>
</tr>
<tr>
<td>14</td>
<td>Exercise time to exhaustion</td>
<td>25</td>
<td>Houston et al. 1979</td>
</tr>
<tr>
<td>28</td>
<td>Exercise time to exhaustion</td>
<td>7.6</td>
<td>Coyle et al. 1986</td>
</tr>
<tr>
<td>28</td>
<td>Exercise time to exhaustion</td>
<td>21</td>
<td>Madsen et al. 1993</td>
</tr>
<tr>
<td>=30</td>
<td>100, 200 m swim</td>
<td>3–13</td>
<td>Mujika et al. 1995</td>
</tr>
<tr>
<td>32</td>
<td>Peak power output</td>
<td>18.2</td>
<td>Nichols et al. 2000</td>
</tr>
<tr>
<td>35</td>
<td>Exercise time to exhaustion</td>
<td>23.8</td>
<td>Fardy 1969</td>
</tr>
<tr>
<td>60</td>
<td>Peak power output</td>
<td>3.5–8.8</td>
<td>Giada et al. 1998</td>
</tr>
</tbody>
</table>
of training cessation in 11 college soccer players (Fardy 1969), and endurance-trained female runners’ oxygen uptake during a standardized submaximal exercise task increased significantly by about 3–8% after 12 weeks of training cessation (Drinkwater & Horvath 1972). In line with these results, Coyle et al. (1985) observed that a submaximal exercise bout requiring a $\dot{V}O_2$ of $3.11 \pm 0.23$ L·min$^{-1}$ ($74 \pm 2\%$ of $\dot{V}O_{2\text{max}}$) when their athletes were trained, elicited a $\dot{V}O_2$ of $3.20 \pm 0.25$ L·min$^{-1}$ ($90 \pm 3\%$ of $\dot{V}O_{2\text{max}}$) after 84 days without training. Compared with values obtained at the peak of their training season, amateur cyclists aged 24 ± 6 and 55 ± 5 years have been shown to decrease their peak power output by, respectively, 3.5 and 8.8% during 2 months without training (Giada et al. 1998). Finally, an injured female masters athlete who could not train for 32 days, decreased her peak power output by 18.2% and her muscular resistance to fatigue measured by a timed ride to exhaustion at 110% of peak power output by 16.6%, as compared with values assessed when she was at the peak of her competitive season, 2 days before getting injured (Nichols et al. 2000).

Training cessation and strength performance

According to the available body of data, athletes can maintain or suffer a limited decay in their muscular strength during short periods of training stoppage. Fourteen days of training cessation did not significantly change 12 weightlifters’ one repetition maximum bench press (−1.7%) and squat (−0.9%) performance, isometric (−7%) and isokinetic (−2.3%) concentric knee extension force, and vertical jump (1.2%) values. On the other hand, isokinetic eccentric knee extension force and surface electromyograph (EMG) activity of the vastus lateralis decreased by 12 and 8.4–12.7%, respectively. These results lead to the conclusion that eccentric strength was specifically affected in strength athletes inactive for a short period of time, but that other aspects of neuromuscular performance were unaltered (Hortobágyi et al. 1993). College swimmers have been shown to maintain their muscular strength measured on a swim bench during 4 weeks of training cessation. However, their swim power, indicative of their ability to apply force during swimming, declined by 13.6% (Neufer et al. 1987). Longer periods of training cessation result in more pronounced declines in the strength performance of strength-trained athletes, but this loss is still limited to 7–12% during periods of inactivity ranging from 8 to 12 weeks. Häkkinen et al. (1981) reported 11.6 and 12.0% decreases in squat–lift and leg extension forces, respectively, following 8 weeks of training stoppage. In addition, maximal bilateral and unilateral isometric force decreased by 7.4 and 7.6%, respectively. This was coupled with decreased averaged maximal bilateral (5.6%) and unilateral (12.1%) integrated EMG. The latter change took place within the first 4 weeks of training cessation (Häkkinen & Komi 1983). Results from the same group of investigators have shown that both muscle atrophy and a diminished neural activation are responsible for the decline in maximal force that takes place during 12 weeks of training cessation, because FT and ST fibre areas, muscle mass and maximal integrated electrical activity were shown to diminish with inactivity (Häkkinen et al. 1981, 1985; Häkkinen & Komi 1983). Studying six endurance-trained subjects during 12 weeks of training cessation, Sinacore et al. (1993) observed a significant reduction in the percentage of initial torque after 30 s of recovery following a 1 min bout of fatiguing exercise of the quadriceps femoris. Changes in FTa and FTb fibre distributions paralleled changes in the rate of torque recovery following the 12-week period of training stoppage. The rate of torque recovery after a standard bout of fatiguing exercise was related to $\dot{V}O_{2\text{max}}$ and may have reflected local muscle endurance exercise adaptations.

Retention of training-induced adaptations

Overuse or any other type of injury can keep athletes from performing their habitual exercises and/or from maintaining their usual training

CONSEQUENCES OF DETERMINING 135
intensity level. In view of the negative effects of physiological characteristics and performance criteria induced by an insufficient training stimulus, any strategy aiming to avoid or reduce detraining would seem worthwhile for the injured or the less active athlete. These strategies, which generally include performing either a reduced training programme or an alternative form of training (i.e. to cross-train), are briefly summarized below.

Reduced training

Reduced training has recently been defined as a non-progressive, standardized reduction in the quantity of training, which may maintain or even improve many of the positive physiological and performance adaptations gained with training (Mujika & Padilla 2000a). From a cardiorespiratory viewpoint, this strategy has been shown to be a valuable procedure to retain many of the training-induced adaptations for at least 4 weeks in highly trained athletes (Neufer et al. 1987; Houmard et al. 1989, 1990a, 1990b; McConell et al. 1993; Martin et al. 1994; Ciuti et al. 1996), and even longer in moderately trained individuals (Hickson & Rosenkoetter 1981; Hickson et al. 1982, 1985; Houmard et al. 1996). As a matter of fact, during periods of reduced training not forced by injury, highly trained athletes have been shown to not change their $\dot{V}O_{2max}$ (Neufer et al. 1987; Houmard et al. 1989, 1990a, 1990b; Madsen et al. 1993; Martin et al. 1994); resting (McConell et al. 1993), submaximal (Houmard et al. 1989, 1990b; McConell et al. 1993) and maximal (Houmard et al. 1989; Madsen et al. 1993) heart rates; submaximal (Houmard et al. 1989; McConell et al. 1993) and maximal (Madsen et al. 1993; McConell et al. 1993) exercise ventilatory volumes; and exercise time to exhaustion (Houmard et al. 1989, 1990b; McConell et al. 1993; Martin et al. 1994). Specific athletic performance, on the other hand, can decline rapidly in highly trained athletes despite the use of reduced training strategies (Neufer et al. 1987; Madsen et al. 1993; McConell et al. 1993).

From a metabolic perspective, it has been shown that submaximal exercise (65, 85 and 95% $\dot{V}O_{2max}$) RER can increase slightly during periods of reduced training (Houmard et al. 1990b; McConell et al. 1993). Exercise blood lactate concentration has been reported not to change (Houmard et al. 1989, 1990b) or to increase (Neufer et al. 1987; McConell et al. 1993). Unchanged insulin action and GLUT-4 concentrations have also been reported as a result of reduced training in recently trained individuals (Houmard et al. 1996).

At the muscle level, research indicates that athletes can readily maintain their lean body mass (Houmard et al. 1989, 1990a, 1990b; McConell et al. 1993), oxidative enzyme activities (Houmard et al. 1990b) and muscular strength (Neufer et al. 1987; Houmard et al. 1990b; Tucci et al. 1992; Martin et al. 1994) by means of reduced training programmes. The retention of these characteristics may be related to a stable hormonal milieu, as suggested by the unchanged testosterone, cortisol and testosterone : cortisol ratio that have been observed in trained distance runners during 3 weeks of reduced training (Houmard et al. 1990a).

A growing body of data in the exercise and sports science literature indicate that maintaining training intensity during periods of reduced training and tapering is of paramount importance for athletes in order to retain training-induced physiological and performance adaptations (Van Handel et al. 1988; Neufer 1989; Houmard 1991; Shepley et al. 1992; McConell et al. 1993; Houmard & Johns 1994; Mujika 1998; Mujika et al. 2000, 2002). Training volume, on the other hand, can be reduced to a great extent without risking detraining effects. Indeed, it has been indicated that this reduction can reach 60–90% of the previous weekly volume, depending on the duration of the reduced training period (Costill et al. 1985, 1991; Cavanaugh & Musch 1989; Neufer 1989; Houmard et al. 1990a, 1990b, 1994; Houmard 1991; D’Acquisto et al. 1992; Johns et al. 1992; Shepley et al. 1992; McConell et al. 1993; Gibala et al. 1994; Houmard & Johns 1994; Martin et al. 1994; Mujika et al. 1995, 1996, 2000, 2002; Mujika 1998). Reports from the literature also indicate that if training-induced physiological and performance adaptations are to be maintained by highly trained
athletes during periods of reduced training frequencies, these reductions should be moderate, not exceeding 20–30% (Neufer et al. 1987; Houmard et al. 1989; Neufer 1989; Houmard 1991; Houmard & Johns 1994; Mujika 1998; Mujika et al., 2002).

For additional information on the physiological and performance consequences of periods of reduced training stimulus, readers can consult the reviews by Houmard (1991), Houmard and Johns (1994), Mujika (1998) and Neufer (1989).

Cross-training

It has been suggested that cross-training, defined here as the participation in an alternative training mode exclusive to the one normally used (Loy et al. 1995), could be a useful means to avoid or limit detraining during recovery from a sport-specific injury. The limited body of data available in the literature on the effects of cross-training as opposed to training cessation suggest that whereas moderately trained individuals may maintain fitness and delay deconditioning by performing dissimilar training modes (Moroz & Houston 1987; Claude & Sharp 1991), similar-mode cross-training would be necessary in more highly trained individuals (Loy et al. 1995).

Thorough information on the possible benefits and practical use of cross-training in sports can be found in a review by Loy et al. (1995).

Cross-transfer effect

During prolonged periods of inactivity, neural (increased motor unit synchronization and activation) and muscular (hypertrophy, increased content of creatine phosphate and glycogen) adaptations induced by strength training (Häkkinen et al. 1985) may be in jeopardy. Interestingly, however, a cross-transfer effect (also referred to as cross-education and cross-training) of training-induced strength gains between ipsilateral (i.e. trained limb) and contralateral (i.e. untrained limb) limbs has been repeatedly described in the literature (Hellenbrandt et al. 1947; Coleman 1969; Shaver 1975; Krotkiewski et al. 1979; Moritani & deVries 1979; Houston et al. 1983; Narici et al. 1989; Kannus et al. 1992a; Housh & Housh 1993; Weir et al. 1994; Housh et al. 1996). This phenomenon has obvious implications in limiting muscular detraining during periods of unilateral casting, rehabilitation from injuries or following joint surgery.

Conclusions

Sports injuries imply periods of training cessation or a marked reduction in the habitual physical activity level of athletes, and an insufficient training stimulus brings about a partial or complete loss of previously acquired physiological and performance adaptations, i.e. detraining. Cardiorespiratory consequences of training cessation in athletes include a rapid \( V_{O_2\text{max}} \) decline, though it usually remains above values of sedentary controls. The \( V_{O_2\text{max}} \) loss results from an almost immediate reduction in total blood and plasma volumes, the latter being caused by a reduced plasma protein content. Exercise heart rate increases during maximal and submaximal intensities, but not sufficiently to counterbalance the reduced stroke volume. Therefore, maximal and submaximal cardiac output drops, whereas it may increase at rest. Cardiac dimensions, including ventricular volumes and wall thickness, often decrease. Blood pressure and total peripheral resistance, on the other hand, increase, and ventilatory efficiency is most usually impaired after short periods of training cessation.

A shift towards a higher reliance on carbohydrate as a fuel for exercising muscles is a primary consequence of training cessation from a metabolic standpoint. Even short-term insufficient training results in an increased respiratory exchange ratio at maximal and submaximal exercise intensities. Glucose tolerance and whole-body glucose uptake are rapidly and markedly reduced, due to a decline in insulin sensitivity coupled with a reduced muscle GLUT-4 transporter protein content. Muscle lipoprotein lipase activity decreases while it increases at the adipose tissue level, thus favouring the storage of adipose tissue. In addition, the training-induced antiatherogenic
lipoprotein profile is reversed. Blood lactate concentration increases at submaximal exercise intensities, and the lactate threshold is apparent at a lower percentage of \( \dot{V}O_{2\text{max}} \). These changes, coupled with a base deficit, result in a higher postexercise acidosis. Trained muscle’s glycogen concentration suffers a rapid decline, reverting to sedentary values within a few weeks of training cessation.

Skeletal muscle tissue is characterized by its dynamic nature and extraordinary plasticity, which allows it to adapt to variable levels of functional demands. When these demands are insufficient to retain training-induced adaptations, muscular detraining occurs. This implies alterations in both muscle structure and muscle function. Muscle capillary density could decrease in athletes within 2–3 weeks of training cessation, but this possibility has not been clearly established. Muscle fibre distribution remains unchanged during the initial weeks of training cessation, but there may be a decreased proportion of ST fibres and a large shift from FTa to FTb fibres in endurance athletes, and an increased oxidative fibre population in strength-trained athletes within 8 weeks of training stoppage. A general decline in muscle fibre CSA is rapidly measurable during training cessation in strength and sprint-orientated athletes, whereas fibre area may increase slightly in endurance athletes. The arteriovenous oxygen difference, unchanged following 3 weeks without training, declines after 8 weeks of continued inactivity. Myoglobin concentration, on the other hand, does not seem to be affected by training cessation. Rapid and progressive reductions in oxidative enzyme activities result in a reduced mitochondrial ATP production. These, along with the reduced arteriovenous oxygen difference, are directly related to the decline in \( \dot{V}O_{2\text{max}} \) observed in athletes undergoing long-term training cessation. These muscular characteristics remain, nevertheless, above sedentary values in the detrained athlete. Glycolytic enzyme activities show non-systematic changes during periods of training cessation.

The general loss in cardiorespiratory fitness, metabolic efficiency and muscle respiratory capacity results in a rapid decline in the trained athletes’ endurance performance. This has been shown by impaired performance measures in all-out swims, peak power output and time to exhaustion tests. In contrast, force production declines slowly and in relation to decreased EMG activity. Strength performance in general is thus readily retained for up to 4 weeks of training cessation, but highly trained athletes’ eccentric force and sport-specific power may suffer significant declines within this timeframe.

Reduced training strategies have been shown to delay the onset of detraining at the cardiorespiratory, metabolic and muscular levels in highly trained athletes. A maintenance of training intensity appears to be the key factor for an athlete to retain training-induced physiological and performance adaptations during periods of reduced training, whereas training volume can be reduced by as much as 60–90%. Training frequency reductions, on the other hand, should be more moderate, not exceeding 20–30%.

Performing alternative training modes exclusive to the one normally used by an athlete (i.e. cross-training) may delay detraining, as long as similar-mode exercises are performed. However, even dissimilar-mode cross-training may be beneficial to the moderately trained subject.

Finally, given that a cross-transfer effect between ipsilateral and contralateral limbs is often observed during periods of unilateral strength training, exercising the healthy limb should be recommended during periods of unilateral casting, rehabilitation from injuries or following joint surgery.

References


Giannattasio, C., Seravalle, G., Cattaneo, B.M. et al. (1992) Effect of detraining on the cardiopulmonary...
reflex in professional runners and hammer throwers. 
*American Journal of Cardiology* 69, 677–680.

effects of tapering on strength performance in trained 
athletes. *International Journal of Sports Medicine* 15, 
492–497.

Gollnick, P.D., Moore, R.L., Riedy, M. & Quintinskie Jr, 
J.J. (1984) Significance of skeletal muscle oxidative 
enzyme changes with endurance training and detraining. 

fiber and motor unit types. *Exercise and Sport Sciences 
Reviews* 21, 331–362.

Häkkinen, K. & Alen, M. (1986) Physiological perform-
ance, serum hormones, enzymes and lipids of an 
elite power athlete during training with and without 
androgens and during prolonged detraining. A case 

Häkkinen, K. & Komi, P.V. (1983) Electromyographic 
changes during strength training and detraining. 

of combined concentric and eccentric strength train-
ing and detraining on force-time, muscle fiber and 
metabolic characteristics of leg extensor muscles. 

Häkkinen, K., Alén, M. & Komi, P.V. (1985) Changes 
in isometric force- and relaxation-time, electromyo-
graphic and muscle fibre characteristics of human 
skeletal muscle during strength training and detraining. 

Postprandial lipemia in endurance-trained people 
during a short interruption to training. *Journal of 
Applied Physiology* 84, 1895–1901.

Hawley, J.A. (1987) Physiological responses to detraining 

and Nutritional Strategies for Sport. Allen & Unwin, 
St Leonards, NSW, Australia.

Heath, G.W., Gavin III, J.R., Hindlerliter, J.M., Hagberg, 
of exercise and lack of exercise on glucose tolerance 

Cross education: the influence of unilateral exercise 

training frequencies and maintenance of increased 
aerobic power. *Medicine and Science in Sports and 
Exercise* 13, 13–16.

Hickson, R.C., Kanakis Jr, C., Davis, J.R., Moore, A.M. 
on aerobic power, endurance, and cardiac growth. 

Hickson, R.C., Foster, C., Pollock, M.L., Galassi, T.M. & 
Rich, S. (1985) Reduced training intensities and loss of 
aerobic power, endurance, and cardiac growth. 

changes in skeletal muscle. *International Journal of 
Sports Medicine* 7, 187–204.

Hortobágyi, T., Houmard, J.A., Stevenson, J.R., Fraser, 
detraining on power athletes. *Medicine and Science in 
Sports and Exercise* 25, 929–935.

Houmard, J.A. (1991) Impact of reduced training on 


Houmard, J.A., Kirwan, J.P., Flynn, M.G. & Mitchell, 
J.B. (1989) Effects of reduced training on submaximal 
and maximal running responses. *International Journal of 
Sports Medicine* 10, 30–33.

Houmard, J.A., Costill, D.L., Mitchell, J.B., Park, S.H., 
Fink, W.J. & Burns, J.M. (1990a) Testosterone, cortisol, 
and creatine kinase levels in male distance runners 
during reduced training. *International Journal of Sports 
Medicine* 11, 41–45.

Houmard, J.A., Costill, D.L., Mitchell, J.B., Park, S.H., 
Hickner, R.C. & Roemmich, J.N. (1990b) Reduced 
training maintains performance in distance runners. 

Effect of short-term training cessation on perform-

Training cessation does not alter GLUT-4 protein 

Houmard, J.A., Scott, B.K., Justice, C.L. & Chenier, T.C. 
(1994) The effects of taper on performance in distance 

Effect of reduced training and training cessation on 

Housh, D.J. & Housh, T.J. (1993) The effect of unil-
ateral velocity-specific concentric strength training. 
*Journal of Orthopaedic and Sports Physical Therapy* 17, 
252–256.

effects of eccentric-only resistance training and 
detraining. *International Journal of Sports Medicine* 17, 
145–148.


Chapter 7

Physiological and Functional Implications of Injury

CHRISTOPHER J. STANDAERT AND STANLEY A. HERRING

Introduction

As even the casual reader of this text can discern, musculoskeletal injury in the athlete can involve a wide range of anatomical structures and physiological variables. Understanding the ways in which injury and treatment affect the athlete’s physical performance is central to developing both an acute treatment strategy and a long-term management plan for the athlete. Injury clearly can involve significant tissue damage, resulting in pain and reduced physical performance. However, injury needs to be viewed in the setting of the entire athlete, not just the local area of acute tissue damage. Prior injuries, strength or flexibility imbalance, skill, technique, or any of a number of other factors affecting the full functional kinetic chain can have a strong bearing on how to approach rehabilitation in a given individual.

As described by Kibler (1998), the term ‘kinetic chain’ refers to the sequencing of individual body segments or joints that must be moved in a specific order and manner to allow for the efficient accomplishment of a given task. The task of pitching a baseball is an example of one such task (Fig. 7.1). This skill, along with most of those in the throwing or hitting sports, actually involves two separate chains of motion: one extends from the ground and planting leg through the hip and trunk to the landing leg; the other extends from the ground and planting leg through the pelvis and trunk to the shoulder, elbow, wrist and, ultimately, the hand (Kibler 1998). Injuries or deficits anywhere along these chains can decrease the efficiency of the task and cause overload on other areas with resultant dysfunction or tissue failure. In considering the local and systemic effects of injury and inactivity, it becomes apparent that any element of the kinetic chain can either be a possible causative factor in injury or be at risk for the deleterious effects of injury. In the case of a baseball pitcher with a sore shoulder, for example, the local tissue injury in the shoulder may be partly related to relative muscular imbalance in the shoulder or inflexibility in the lumbar spine. Resting the pitcher with the intent of allowing the

Fig. 7.1 Illustration of the kinetic chains involved in pitching a baseball. The first chain extends from the ground and planting leg to the hip, trunk and down the landing leg. The second chain runs from the ground and planting leg to the hip and trunk then to the shoulder, elbow, wrist and hand of the throwing arm. (Adapted from Kibler 1998.)
Consequences of injury may result in numerous less desirable consequences. These might include reductions in endurance, strength, mobility or neuromuscular control, all of which are necessary for the production of appropriate arm speed, positioning and release necessary to achieve optimal ball speed and movement. Awareness of the potential for these deficits is crucial in the rehabilitation of the athlete. This chapter will present a format for understanding injury in this context and include a discussion of the physiological and functional implications of injury as they relate to rehabilitation.

Rehabilitation

Sports injuries are an extremely common event for athletes (see Chapter 1). Reports of injury rates and severity vary widely by sport and also by methods of measurement (Henry et al. 1982; Keller et al. 1987; Saal 1991; Bennell & Crossley 1996; Murtaugh 2001). None the less, reported injury rates can be extremely high, including a 12.7% incidence of injury in adolescent wrestlers during participation in a single tournament (Lorish et al. 1992), a greater than 50% injury rate for high school football players for a single season (DeLee & Farney 1992), a 69% injury rate for a professional basketball team over 7 years (Henry et al. 1982) and a lifetime incidence of >70% for interfering shoulder pain in elite swimmers (McMaster & Troup 1993). The natural history of soft tissue injuries is not necessarily benign, either—a recent study on ankle sprains in the general population (the vast majority of which were not treated with extensive rehabilitation) reported that 72.6% of subjects had persisting symptoms 6–18 months after injury (Braun 1999). An initial injury may also have a significant impact on the risk of further injury. A review on injuries in soccer players noted that prior injury was one of the factors most strongly associated with new injury and that 20% of ‘minor’ injuries were followed by a more severe injury within 2 months (Keller et al. 1987). Macera (1992) also noted that there is a strong effect of previous injury on the risk of further injury in runners. The precise reasons for the increased risk of reinjury are not entirely clear, but several possibilities exist, including incomplete recovery from the initial injury, incomplete rehabilitation, uncorrected biomechanical problems, or physiological, anthropomorphical or technique issues related to an individual athlete (Keller et al. 1987; Rutherford et al. 1990; Macera 1992; Taylor et al. 1993; Jonhagen et al. 1994). Clearly, athletic injuries have a significant impact on a large number of athletes, teams and sporting events. Interventions that may lessen the ultimate impact of an injury upon an athlete’s performance, including comprehensive rehabilitation, may have a dramatic effect not just for the individual athlete, but for sports competition as a whole.

The rehabilitation plan for an individual athlete is based upon an assessment of the athlete’s specific injury and more global functional properties, along with an understanding of the functional consequences of the injury and treatment. First, the actual tissue injury must be diagnosed accurately. The presentation of the injury, or the clinical symptom complex (Herring 1990), needs to be identified, as do the type of injury and the physiological and biomechanical factors associated with its occurrence. Finally, the acute method of treatment needs to be established with a thorough knowledge of the local or systemic effects of the chosen modality (e.g. immobilization, surgery, etc.). From this base, an appropriate rehabilitation strategy can be developed that addresses the full functional capacity of the athlete (Herring & Kibler 1998).

Although injury is often associated with focal tissue damage, the goal of treatment for an injured athlete extends far beyond achieving tissue homeostasis. The athlete’s goal is to return to sport at the highest level of performance possible. This means that medical professionals have to deal with all aspects of the athlete’s functioning. As will be discussed, injury and the subsequent acute treatment can result in diminished soft tissue tensile strength, decreased aerobic capacity, cellular and neurological changes, and residual disability even in the absence of symptoms. Rehabilitation needs to address all of these...
issues, with an initial goal of symptom resolution and an ultimate goal of restored function and continued fitness to minimize the risk of recurrent injury and preserve long-term performance (Herring & Kibler 1998). This concept is important in understanding why rehabilitation is necessary in the first place and why ‘rehabilitation’ really refers to an ongoing process of training and maintaining the injured athlete.

**Injury**

Specific tissues respond differently to the stress and strain associated with trauma (see Chapters 3–5). Generally speaking, however, injury occurs when the forces applied to a given anatomical structure exceed the physiological tolerance or reparative capabilities of that structure. As a gross step in assessing injuries, they can be broken down into two major types: macrotraumatic or microtraumatic (Leadbetter 1994; Quillen et al. 1996; Herring & Kibler 1998). Macrotrauma refers to acute, extreme forces that overwhelm the tensile or structural properties of a given tissue. This is generally thought of as resulting in the acute deformation or disruption of previously normal structures, such as might occur with an acute ankle fracture or anterior cruciate ligament tear. Microtrauma, on the other hand, refers to smaller forces that cause lesser degrees of cellular or structural disruption and occur at a rate or severity that exceeds the tissue’s ability to enact appropriate repair or maintenance. This can result in weakening of the structural properties of the affected tissue, adaptations of function in adjacent or compensatory structures, pain, impaired performance and, potentially, tissue failure. There is a continuum of injury present between these two mechanisms of ‘acute’ and ‘chronic’ injury, however; many apparently acute or macrotraumatic events causing disruption are strongly related to underlying degeneration, weakness or misuse resulting from a lengthier microtraumatic process of tissue degeneration. The vast majority of acute tendon ruptures, in fact, are associated with prior pathological change in the tendon (Leadbetter 1994). This conceptual model reinforces the concept that injury occurrence, prevention and treatment represent ongoing, dynamic processes occurring within the body’s structural and cellular balance (Fig. 7.2).

A structural format for approaching the treatment of an athletic injury needs to include a full understanding of the clinical and physiological alterations induced by or associated with injury (Table 7.1). The injured athlete is typically being assessed by medical personnel for complaints of pain, weakness, oedema or dysfunction. The presenting symptoms can be referred to as the clinical symptom complex. For a patient with a rotator cuff injury, there may be local complaints of pain, weakness and diminished range of motion.

![Fig. 7.2 The injury cycle and feedback mechanisms, illustrating the multifactorial nature of athletic injury. (From Kibler et al. 1992, with permission.)](image-url)
all affecting the injured shoulder. The actual tissue injury represents the tissue injury complex. In the above-mentioned athlete with a rotator cuff injury, this may be a partial tear of the supraspinatus tendon. Frequently, medical assessment focuses on these two components of the injured athlete. The actual symptomatic tissue injury, however, may be accompanied by problems with associated structures that have been overloaded, suboptimally lengthened or otherwise impaired by the acute injury, or that may be contributing to the injury. These related problems are referred to as the tissue overload complex (Herring 1990; Herring & Kibler 1998). In the athlete with a shoulder injury, this may include tensile load on the posterior shoulder capsule due to dynamic instability of the glenohumeral joint or excessive load on the rhomboids related to relative lateral translation of the scapula.

In looking beyond local tissue injury, the clinician also has to assess underlying biomechanical or technique issues involved with the performance of an injured athlete. Continuing with the shoulder example, studies of relative isokinetic strength in uninjured elite water polo players and collegiate tennis players have shown significant strength imbalance at the shoulder with a relative increase in strength of the shoulder internal rotators compared with the external rotators (McMaster et al. 1991; Chandler et al. 1992). This relative imbalance has been hypothesized to be related to shoulder injuries in those athletes due to an alteration of the normal mechanics of the glenohumeral joint (McMaster et al. 1991; Chandler et al. 1992). This is an example of the functional biomechanical deficit complex which includes abnormalities in strength or flexibility associated with injury. Finally, alterations in mechanics, motion or task performance may result from injury or functional biomechanical deficits. These associated problems are referred to as the subclinical adaptation complex. As this name would imply, these adaptations may not be directly associated with symptoms but can contribute to progression or perpetuation of injury and impaired performance (Herring 1990; Kibler et al. 1991; Herring & Kibler 1998). An example of this would be a baseball pitcher who lowers the elbow of his or her throwing arm and changes the release point of the ball due to a combination of shoulder pain, dynamic instability and internal rotation inflexibility.

Table 7.1 Complete diagnosis of rotator cuff tendonitis/impingement in a baseball pitcher. (Adapted from Herring 1990.)

<table>
<thead>
<tr>
<th>Clinical symptom complex</th>
<th>Impingement symptom pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior–superior glenohumeral instability</td>
<td></td>
</tr>
<tr>
<td>Decreased pitch velocity</td>
<td></td>
</tr>
<tr>
<td>Prolonged recovery time</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tissue injury complex</th>
<th>Tear in posterior capsule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tear in rotator cuff</td>
<td></td>
</tr>
<tr>
<td>Glenoid labral attrition</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tissue overload complex</th>
<th>Tensile load on posterior shoulder capsule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tensile load on posterior shoulder muscles</td>
<td></td>
</tr>
<tr>
<td>Tensile load on scapular stabilizing muscles</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Functional biomechanical deficit complex</th>
<th>Internal rotation inflexibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>External/internal rotation strength imbalance</td>
<td></td>
</tr>
<tr>
<td>Lateral scapular slide</td>
<td></td>
</tr>
<tr>
<td>Alteration of trunk motion with increased passive lumbar lordosis and decreased lateral lumbar flexion</td>
<td></td>
</tr>
<tr>
<td>Decreased hip mobility</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclinical adaptation complex</th>
<th>Altered shoulder or elbow positioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decreased pitch velocity</td>
<td></td>
</tr>
</tbody>
</table>
Viewing injury in the context just described allows medical personnel to fully evaluate all of the contributing biomechanical and physiological factors affecting an athlete’s injury, impaired performance and full symptom complex. Once this is assessed, an appropriate strategy for correcting deficits in all relevant areas can be addressed, and an acute injury treatment plan can advance to a full management programme for the athlete.

**Tissue damage—the tissue injury complex**

Although the general inflammatory cascade is similar in all acute soft tissue injuries, the distinct structural and reparative properties of different tissues impact on how this process affects recovery and performance. The injury and repair processes for specific tissues are well addressed in other chapters in this book (see Chapters 3–5). However, some specific aspects of soft tissue injury and repair will be briefly reviewed in order to more fully develop the underpinnings of rehabilitation treatment.

Ligament injuries frequently occur as a result of acute macrotraumatic forces applied to the ligament causing structural disruption of varying degrees. The amount of force necessary to cause ligamentous disruption is dependent upon the size of the ligament, the age of the individual and the joint position at the time of injury. The healing process results in scar formation with collagen deposition but is frequently inadequate to withstand future load bearing (Frank 1996). Even in the best of circumstances, maximal recovery may take more than 1 year and still result in a decrement of 30–50% in tensile strength (Leadbetter 1994). Immobilization may additionally impair the structural properties of an injured ligament (Frank 1996; Salter 1996). As the functional role of a ligament is to serve as a passive support for joint positioning and to potentially provide proprioceptive feedback, the rehabilitation of ligamentous injuries needs to include work on restoring or improving dynamic joint control through the kinetic chain in order to reinforce the stabilizing properties of an injured ligament. The loss of joint stability associated with an acute ligament injury also needs to be considered during the early phase of rehabilitation in order to avoid further injury to the ligament complex.

Tendon injuries are slightly different from ligamentous injuries and, subsequently, pose different issues in the rehabilitation process. Tendons may also be acutely injured due to macrotraumatic forces applied across the musculotendinous unit. However, tendons are also the structures most frequently involved in overuse, or, perhaps more appropriately termed, chronic overload syndromes, which involve repetitive loading with partial disruption of the tendon structure (Herring & Nilson 1987; Curwin 1996). In overload injuries, the forces applied to an injured tendon may be within a physiological range for the tendon but they occur at a frequency that prohibits adequate recovery or repair processes. This may ultimately lead to failure of the tendon structure (Curwin 1996). Tendon repair mechanisms would imply that progressive loading of the injured structure is necessary to optimize the restoration of the tensile strength. Given the frequency with which extrinsic factors are associated with tendon injuries, it is also important to assess an athlete for issues of flexibility, strength imbalance, equipment and technique in order to minimize deleterious effects that may be associated with external appliances or maladaptive motion patterns. These issues could potentially involve the tissue overload complex, the functional biomechanical deficit complex or the subclinical adaptation complex, and illustrate the importance of recognizing how all of these areas contribute to the problems of an injured athlete.

Muscular injuries are, yet again, distinct from isolated tendinous or ligamentous injuries. Muscle injuries may occur in the midsubstance of the muscle proper, such as in a contusion, or in the region of the myotendinous junction, as in a strain. In considering differences between tendinous and muscular injuries, the region of the musculotendinous junction warrants particular attention. Muscle strain injury is felt to occur during elongation of the muscle, typically during
forceful eccentric contraction, and is particularly common in sports involving sprinting or jumping (Noonan & Garrett 1992; Taylor et al. 1993; Jonhagen et al. 1994; Bennell & Crossley 1996; Garrett 1996; Jarvinen et al. 2000). The vast majority of muscular strain injuries have been shown to occur at the region of the musculotendinous junction, generally affecting the muscle immediately adjacent to the junction rather than in the junction itself (Taylor et al. 1993; Garrett 1996; Malone et al. 1996). ‘Two-joint’ muscles, such as the rectus femoris, gastrocnemius and hamstrings, are particularly vulnerable to this type of injury (Noonan & Garrett 1992; Mair et al. 1996; Jarvinen et al. 2000). The tensile forces generated by the muscle during maximal eccentric loading combined with the forces associated with elongation of the static viscoelastic components of the muscle probably contribute to the frequency with which strain injury occurs during moments of forceful eccentric contraction of the muscle (Malone et al. 1996). There are, additionally, distinctive properties of the sarcomeres and membranous structures adjacent to the musculotendinous junction that may contribute to this area being particularly vulnerable to strain injury (Noonan & Garrett 1992).

Healing of muscular strain injuries, as well as of disruptions of the muscle more distant to the site of the musculotendinous junction, consists of scar deposition with regeneration and ingrowth of contractile elements and nerves (Jarvinen et al. 2000). Acutely, there is a significant decline in contractile strength and peak load tolerance in injured muscle, and the scar is at its weakest point until 10–12 days after injury (Taylor et al. 1993; Jarvinen et al. 2000). Tensile strength recovery also appears to lag behind the recovery of force-generating capacity (tension generation), making the muscle perhaps more vulnerable to repeat injury than it may appear by tension-generating capacity alone (i.e. gross measures of ‘strength’) (Garrett 1996). Scar formation may significantly impair the ability of contractile elements in injured muscle to repair, and contractile ability may be reduced by 10–20% long term (Jarvinen & Lehto 1993; Leadbetter 1994). The type of treatment chosen and the length of immobilization may also have a significant impact on the actual reparative processes within injured muscle and the degree to which contractile elements penetrate the connective tissue (Jarvinen & Lehto 1993).

Accurately assessing the nature and severity of the tissue injury in an affected athlete, i.e. the tissue injury complex, is crucial in determining an appropriate rehabilitation approach to the patient. The type of tissue injured, the degree of structural damage, the concurrent structures affected and the potentially deleterious effects of immobilization on adjacent structures all need to be taken into account in treatment. Although different tissues are discussed in isolation above, an injury frequently results in alterations to several different tissue components. For example, a surgically treated acute tear of the anterior cruciate ligament in an athlete may ultimately result in alterations of bone, tendon and ligament with additional components of muscular atrophy, capsular tightness, aerobic deconditioning and alterations in neural control. Understanding each of the individual components associated with an injury allows for a more comprehensive and precise approach to treatment.

**Immobilization and rest**

Acute injury of soft tissue or bone is frequently managed by a period of immobilization of the injured body part or by rest of the injured area or the athlete as a whole. This is done to allow for adequate healing and restoration of tensile strength so that the injured structure may begin to bear physiological loads with less risk for recurrent injury. Periods of immobilization required for the treatment of different injuries can vary highly. For example, ankle or radial fractures may frequently be immobilized in a rigid cast for 4–6 weeks, acute spinal fractures may be treated with a rigid brace for up to 3 months, and some authors recommend the use of a rigid brace for 6–12 months in the management of spondylolisthesis (Steiner & Micheli 1985; Tropp & Norlin 1995; McDowell 1997; Byl et al.
Although often viewed as a useful tool in the management of musculoskeletal injuries, immobilization and inactivity can have significant deleterious effects on the injured athlete. These changes may affect how an athlete will be able to progress through a rehabilitation programme after injury. Understanding the scope and time course of deficits involved with rest and immobilization allows health care professionals to apply both preventative and restorative care in the course of recovery.

At the muscular level, there are significant changes in the function and structural properties of the musculotendinous unit treated with immobilization. These changes include reductions in strength, limb volume, limb weight and the cross-sectional area of muscles as a whole and of individual muscle fibres (Muller 1970; Wills et al. 1982; Booth 1987; Veldhuizen et al. 1993; Bloomfield 1997; Zarzhevsky et al. 1999). There are, additionally, selective changes in enzymatic function and protein synthesis and alterations to the musculotendinous junction with reduced contact area between the muscle cells and tendineal collagen fibres (Wills et al. 1982; Appell 1990; Kannus et al. 1992; Hortobagyi et al. 2000). All of these factors play a role in the functional implications for managing patients with injuries treated with immobilization.

When looking strictly at muscle strength, the losses associated with immobilization can be profound. A recent study by Hortobagyi et al. (2000) found that healthy volunteers placed in a fibreglass cast for 3 weeks sustained an average 47% decrease in eccentric, concentric and isometric strength of knee extension. Similar decreases in peak isometric torque of 53% for knee extension and 26% for knee flexion were noted by Veldhuizen et al. (1993) after 4 weeks of cast immobilization in healthy volunteers (Fig. 7.3). Geboers et al. (2000) found a 28% decrease in dorsiflexion torque for individuals who underwent cast immobilization for 4–6 weeks after ankle fractures. A study of individuals casted for 8 weeks after open reduction and internal fixation of an ankle fracture found a reduction of about 50% for ankle plantar flexion peak isometric torque with reductions in isokinetic torque for all angular speeds and positions (Shaffer et al. 2000). The strength losses seem to be more dramatic with the use of rigid immobilization than with either limb suspension or bed rest, although these also result in dramatic reductions in strength over a fairly rapid time course (Berg et al. 1991; Bloomfield 1997). This may suggest that there is a beneficial effect on relative strength loss by the preservation of some degree of joint mobility, although the true magnitude of any such effect is not clear currently. Declines in muscular strength with immobilization...
or disuse appear to be most prominent in anti-gravity extensor muscles such as the quadriceps and gastrocnemius/soleus (Appell 1990; Dittmer & Teasell 1993; Veldhuizen et al. 1993; Bloomfield 1997; Hortobagyi et al. 2000). Overall, the most rapid period of strength loss seems to be early in the course of immobilization with little additional loss occurring after the first week (Muller 1970; Wills et al. 1982; Appell 1990).

The losses in strength are paralleled by muscle atrophy and losses in muscle fibre size (Fig. 7.4). Human studies have shown a decrease in fibre size by up to 14–17% after 72 h of immobilization (Lindboe & Platou 1984; Booth 1987). In the study mentioned previously by Veldhuizen et al. (1993), healthy subjects who were placed in a long leg cast for 4 weeks were found to have sustained a 21% loss in quadriceps cross-sectional area by computed tomography and a 16% decrease in fibre diameter.

Although the strength losses associated with immobilization occur rapidly, the recovery of strength and mass may take a very long time, and strength may never recover completely in some cases. Rutherford et al. (1990) found a marked reduction in strength and cross-sectional area of previously injured and immobilized limbs compared with the contralateral, unjured limbs in patients assessed 1–5 years after their injuries. Similarly, other authors report prolonged decrements in strength or muscle mass for weeks to months after immobilization (Grimby et al. 1980; Wills et al. 1982; Appell 1990; Tropp & Norlin 1995; Zarzhevsky et al. 1999; Cruz-Martinez et al. 2000). The rate and extent of recovery may be dependent upon the type and intensity of retraining applied. Both the effects of immobilization and the response to training appear to be speed and task specific (Grimby et al. 1980; Wills et al. 1982; Rutherford 1988; Appell 1990; Zarzhevsky et al. 1999). These issues clearly have a major bearing upon any rehabilitation plan for an injured athlete who has been treated with rest and/or immobilization.

**Neural changes**

Loss in muscle mass may not be the only factor affecting muscular performance following rest or immobilization. There appear to be neural changes that affect relative muscular efficiency. This is suggested by changes in the relative electromyograph (EMG) activity per unit of force produced, a decrease in the number of functioning motor units and a decrease in reflex potentiation (Bloomfield 1997). Cruz-Martinez et al. (2000) found EMG abnormalities felt to be consistent with quadriceps motor neurone inhibition in patients studied after 1–4 months of knee immobilization after injury. Seki et al. (2001a, 2000b) studied the effects of 6 weeks of immobilization on the function of the first dorsal interosseus
muscle of the hand and found that the maximal motor neurone firing rate was decreased transiently. The changes in maximal firing rate were more profoundly affected during the first 3 weeks of immobilization than in the latter 3 weeks and seemed to recover close to normal after 6 weeks following the period of immobilization. The authors felt that changes in the contractile properties of muscle after immobilization can be causally related to the alterations in firing rate of the motor neurones.

In addition to alterations at the spinal cord or muscular level, there is evidence that immobilization affects cortical function as well. A study by Liepert et al. (1995) using transcranial magnetic stimulation identified a decrease in the motor cortex area associated with tibialis anterior function in patients immobilized at the ankle (Fig. 7.5). The cortical changes reversed rapidly after voluntary muscular function, but suggest that cortical reorganization may occur with immobilization. This may potentially play a role in reduced muscular efficiency and motor control following injury and disuse.

An additional factor affecting muscular efficiency after injury may be the influence of degenerative changes or joint effusion on muscular function. It has been proposed that the presence of a joint effusion or degenerative joint changes may cause reflex inhibition of certain muscle groups or fibre types (Spencer et al. 1984; Young et al. 1987; Hurley & Newham 1993; Young 1993; Hopkins et al. 2001). This concept is supported, in part, by work showing differential reflex inhibition or relative excitability of selected muscle groups at a joint affected by an effusion (Young et al. 1987; Hopkins et al. 2001). Spencer et al. (1984) studied the effects of progressive instillation of fluid into the knee joints of healthy volunteers in order to assess the effects of an effusion on reflex inhibition. They found that increasing volumes of fluid instilled into the joint were associated with reductions in the H-reflex amplitude recorded from the quadriceps muscles. The H (or Hoffman) reflex is felt to reflect the relative excitability of the motor neurone pool. The degree of change in the amplitude of the evoked response was related to the volume instilled in a linear manner (Fig. 7.6). The changes were not seen in anaesthetized joints, consistent with a mechanism of motor neurone inhibition mediated by afferent receptors in the knee joint or capsule.

The potential inhibitory effect of a joint effusion on motor neurone function may represent a confounding factor when comparing studies that involve individuals treated with immobilization after an injury (in which there may well be an associated effusion) to those that involve the immobilization of healthy volunteers (who probably do not have an effusion present). However, in clinical practice, immobilized joints are very frequently associated with an injury or surgical intervention that may result in an effusion. Understanding the potential for neurological changes seen with either immobilization alone or with an effusion may allow clinicians to better understand the functional properties of the limb of an injured athlete affected by an effusion, a period of immobilization, or both.
Vascular changes

Along with muscular and neural changes associated with immobilization, there also appear to be significant local vascular changes. Immobilization affects both the vascular density in the muscle and the three-dimensional structure of the vessels themselves (Kvist et al. 1995; Oki et al. 1998). In a study on the rat gastrocnemius, Kvist et al. (1995) found that 3 weeks of immobilization resulted in a 30% reduction in capillary density in the musculotendinous junction of the immobilized muscle. As most of the blood supply to a tendon is felt to come from the muscle vasculature, a decrease in the capillary bed at the musculotendinous junction may be a contributing factor to subsequent injury due to loss of tensile strength in the tendon (Kvist et al. 1995). The potential for recovery of the vascular system appears to be good with remobilization and occurs much more rapidly in muscles mobilized earlier and in those subjected to progressively increased physical training (Jarvinen & Lehto 1993; Kvist et al. 1995).

Joint function

Immobilized joints are subject to multiple potential changes after immobilization, including contracture, adhesions, cartilage atrophy, erosions at areas of bony contact and reduced load tolerance of ligamentous attachments (Akeson et al. 1987; Dittmer & Teasell 1993). Loss of joint motion is a significant concern following immobilization for an injury. Human and animal studies have shown significant deficits in motion of affected joints following immobilization (Akeson et al. 1987; Tropp & Norlin 1995; Reynolds et al. 1996; Salter 1996; Schollmeier et al. 1996; Byl et al. 1999; Trudel et al. 1999; Trudel & Uhthoff 2000). In a rat model using immobilization at the knee, Trudel et al. (1999) found that joint contracture occurred as little as 2 weeks after immobilization and progressed at an average rate of $3.8^\circ$-week$^{-1}$ for 16 weeks with an apparent plateau reached after that point (Fig. 7.7). In a study on the effects of cast immobilization for distal radial fractures treated non-operatively, Byl et al. (1999) found

**Fig. 7.6** Reduction in H-reflex amplitudes in the quadriceps muscles as a function of the volume of saline introduced into the knee joint. (From Spencer et al. 1984, with permission.)
marked reductions in range of motion after cast removal (average cast time = 4.1 weeks). Their results included losses of greater than 50% of expected motion for all planes in the wrist and restriction of pronation and supination to 40% of normal.

Studies on the persistence of joint contracture after immobilization and injury vary, and the actual results in a given case are probably dependent upon multiple issues including the actual nature of the injury and the extent and duration of immobilization. After 4 weeks of immobilization of the knee in a plaster cast in healthy human volunteers, Veldhuizen et al. (1993) found that range of motion returned to normal within 3 days after cast removal. In a canine model of shoulder immobilization, Schollmeier et al. (1996) found that 12 weeks of cast use resulted in significant losses in passive range of motion initially that were normalized by 12 weeks after cast removal. In contrast to this, Tropp and Norlin (1995) found significant declines in range of motion at the ankle 10 weeks after surgical treatment for ankle fractures with motion close to normal by 12 months after injury. Significant joint dysfunction well beyond this timeframe following immobilization has also been recognized in the medical literature for well over a century (Salter 1996).

Beyond loss of range of motion, numerous other aspects of joint function and structure are affected by immobilization. It has been proposed that joint function and homeostasis are dependent upon motion and use (Akeson et al. 1987; Salter 1996). In the absence of these factors (i.e. with immobilization), the functional and structural properties of the joint cannot be maintained. This leads to a number of documented changes in the joint structure. These include the degradation of articular cartilage, adherence of the synovial membrane to the underlying cartilage and within synovial folds, synovial hyperplasia, destruction of ligamentous attachment sites and proliferation of connective tissue into the joint space (Akeson et al. 1987; Salter 1996; Schollmeier et al. 1996). Synovial fluid is also affected, with reductions in hyaluronic acid levels noted following immobilization in an animal model (Pitsillides et al. 1999). There is, additionally, degradation of the collagen fibre structure of associated ligaments, probably presenting a potential source of recurrent injury (Akeson et al. 1987).

**Cardiovascular and aerobic capacity**

In addition to changes in locally involved structures, immobilization and rest can have profound effects on overall cardiovascular and aerobic capacity. Dramatic declines in $\dot{V}O_2$ max and maximal exercise performance can occur with extended rest (Neufer 1989; Convertino 1997). The rate of loss of aerobic capacity has been estimated at 0.9% per day for 30 days of bed rest (Convertino 1997). The decline in $\dot{V}O_2$ max initially appears to be due to a reduction in circulating blood volume and venous return with a subsequent decline in cardiac output and stroke volume (Neufer 1989;
Peripheral mechanisms of decreased muscle capillarization, decreased vascular conductance and decreased muscular oxidative enzyme capacity also occur with detraining but do not seem to play a major role in initial declines in $V_o_{2\text{max}}$. These changes may be more important in submaximal exercise tolerance, however (Neufer 1989). Changes in $V_o_{2\text{max}}$ may be more profound in conditioned athletes, in part due to their greater relative level of aerobic fitness compared to untrained individuals (Convertino 1997).

Losses in aerobic fitness are related to the relative duration, frequency and intensity of training levels that are retained during periods of reduced activity that may follow an injury. $V_o_{2\text{max}}$ can be maintained for up to 15 weeks with reductions in training frequency of up to 66% if the training duration is held constant (Neufer 1989). A combined reduction of training frequency and duration that results in a 70% reduction in energy demand can maintain $V_o_{2\text{max}}$ over a period of 4 weeks (Neufer 1989). Maintaining the intensity of exercise training during periods of reduced frequency or duration of activity, however, is crucial to maintaining $V_o_{2\text{max}}$ (Neufer 1989). In the rehabilitation of an injured athlete, attempts to modify exercise patterns, where appropriate, to allow for continued limited training sessions with maintained cardiovascular intensity may be helpful in the acute stage and to potentially reduce the systemic cardiovascular effects of rest for a local tissue injury.

When viewing the changes that occur in aerobic capacity with extended rest or reduced activity in conjunction with the effects of rest and immobilization on joint function, and vascular, neural and muscular functioning, the need to assess the athlete as a whole becomes more apparent. Local tissue injury combined with the wide-ranging effects frequently associated with the required treatment can present significant challenges to obtaining optimal recovery and performance following an athletic injury. Thoroughly understanding these issues and following the approach to rehabilitation outlined previously (Table 7.1) allows a clinician to incorporate a global view of athletic function into a rehabilitation plan from the onset of injury through resumption of play.

The kinetic chain and muscular balance

As noted previously, comprehensive rehabilitation of an injured athlete requires that a clinician assess the entire kinetic chain for problems that may arise from or contribute to injury (Fig. 7.1). Two of the major issues that relate to the mechanics of, and force transmission through, the kinetic chain are relative strength and flexibility across regions of the chain. Determining the optimal degree of flexibility or muscular balance for a given athlete may not necessarily follow uniform parameters, however, and these factors need to be assessed in the context of a given individual and the demands of his or her sport.

The role of abnormalities in relative muscular balance in the occurrence of athletic injuries has been debated in the literature. A review on the topic by Grace (1985) noted that, although muscular imbalance as a causative factor in athletic injury seemed logical, there was little definitive proof of this. Part of the problem arises from the lack of any clear definition of ‘imbalance’ for a given muscular group. However, there are a number of reports since the date of that review that have identified relative muscular imbalance in studies of athletes postinjury, in screening of uninjured athletes commonly at risk for particular injuries, and as a significant risk factor for injury in prospective studies (Fleck & Falkel 1986; Taimela et al. 1990; Kibler et al. 1991; Knapik et al. 1991; McMaster et al. 1991; Chandler et al. 1992; Jonhagen et al. 1994; Baumhauer et al. 1995; Orchard et al. 1997). Kibler et al. (1991) identified relative ankle inflexibility and deficits in peak plantar flexion torque in the affected legs of runners with plantar fasciitis. A prospective study on the risk of ankle injury in college athletes found that injured and uninjured athletes differed in relative dorsiflexion to plantar flexion and eversion to inversion ratios as measured with isokinetic testing (Baumhauer et al. 1995). Another
prospective study on college athletes by Knapik et al. (1991) also found that muscular imbalance either across the knee or between sides in a given athlete or imbalance in hip motion between sides was associated with a higher risk of injury. Similarly, Orchard et al. (1997) identified low hamstring to quadriceps peak torque ratios as a risk factor for hamstring injury in a prospective study of professional Australian rules footballers. Muscular imbalance about the shoulder has also been described in a variety of ‘overhead’ athletes, such as tennis players, swimmers and water polo players (Fleck & Falkel 1986; McMaster et al. 1991; Chandler et al. 1992). As with muscle imbalance, the role of flexibility in sports injuries remains debated. Issues of flexibility are frequently cited as relating to athletic injury (Keller et al. 1987; Taimela et al. 1990; Taylor et al. 1990; Kibler et al. 1991; Knapik et al. 1991; Jonhagen et al. 1994; Worrell 1994; Baumhauer et al. 1995; Stocker et al. 1995; Bennell & Crossley 1996; Gleim & McHugh 1997). Several studies have found a positive association between injury and greater overall flexibility or ‘ligamentous laxity’, although other studies have failed to show an association between the two (Keller et al. 1987; Taimela et al. 1990; Baumhauer et al. 1995; Stocker et al. 1995; Bennell & Crossley 1996). Some authors note that relative inflexibility of a particular muscular group may be a predisposing factor to injury, particularly musculotendinous strain (Keller et al. 1987; Knapik et al. 1991; Jonhagen et al. 1994; Worrell 1994). Studies by Kibler et al. (1991) and Jonhagen et al. (1994) have shown relative decreases in passive range of motion measures compared with controls in athletes with either plantar fasciitis or prior hamstring strains, respectively. In a review of the role of flexibility in sports injury, Gleim and McHugh (1997) noted that, as with muscular imbalance, there is a lack of any standard definition for ‘flexibility’ and that there are a variety of ways in which flexibility can be assessed (such as static vs dynamic and stretch tolerance vs muscle stiffness). They also noted that specific flexibility patterns are associated with specific sports or even positions within a given sport. In some instances, lesser degrees of flexibility may, in fact, be beneficial in force production associated with rapid muscular contraction (Gleim & McHugh 1997). It is likely that a variety of physiological and anthropomorphic factors allow a given athlete to obtain an elite level of performance in a given task. Traits that are beneficial for some sport-specific tasks may actually be detrimental to the performance of others. The function of an athlete needs to be viewed in the context of that athlete’s particular sport and the entire motion chain that allows for the performance of specific skills. Injury and rehabilitation must encompass this view, and deficits in flexibility (or the presence of excessive laxity) need to be assessed in the performance of a given task.

Studies by Kibler et al. (1991) and Jonhagen et al. (1994) have shown relative decreases in passive range of motion measures compared with controls in athletes with either plantar fasciitis or prior hamstring strains, respectively. In a review of the role of flexibility in sports injury, Gleim and McHugh (1997) noted that, as with muscular imbalance, there is a lack of any standard definition for ‘flexibility’ and that there are a variety of ways in which flexibility can be assessed (such as static vs dynamic and stretch tolerance vs muscle stiffness). They also noted that specific flexibility patterns are associated with specific sports or even positions within a given sport. In some instances, lesser degrees of flexibility may, in fact, be beneficial in force production associated with rapid muscular contraction (Gleim & McHugh 1997). It is likely that a variety of physiological and anthropomorphic factors allow a given athlete to obtain an elite level of performance in a given task. Traits that are beneficial for some sport-specific tasks may actually be detrimental to the performance of others. The function of an athlete needs to be viewed in the context of that athlete’s particular sport and the entire motion chain that allows for the performance of specific skills. Injury and rehabilitation must encompass this view, and deficits in flexibility (or the presence of excessive laxity) need to be assessed in the performance of a given task.

As with muscle imbalance, the role of flexibility in sports injuries remains debated. Issues of flexibility are frequently cited as relating to athletic injury (Keller et al. 1987; Taimela et al. 1990; Taylor et al. 1990; Kibler et al. 1991; Knapik et al. 1991; Jonhagen et al. 1994; Worrell 1994; Baumhauer et al. 1995; Stocker et al. 1995; Bennell & Crossley 1996; Gleim & McHugh 1997). Several studies have found a positive association between injury and greater overall flexibility or ‘ligamentous laxity’, although other studies have failed to show an association between the two (Keller et al. 1987; Taimela et al. 1990; Baumhauer et al. 1995; Stocker et al. 1995; Bennell & Crossley 1996). Some authors note that relative inflexibility of a particular muscular group may be a predisposing factor to injury, particularly musculotendinous strain (Keller et al. 1987; Knapik et al. 1991; Jonhagen et al. 1994; Worrell 1994). Studies by Kibler et al. (1991) and Jonhagen et al. (1994) have shown relative decreases in passive range of motion measures compared with controls in athletes with either plantar fasciitis or prior hamstring strains, respectively. In a review of the role of flexibility in sports injury, Gleim and McHugh (1997) noted that, as with muscular imbalance, there is a lack of any standard definition for ‘flexibility’ and that there are a variety of ways in which flexibility can be assessed (such as static vs dynamic and stretch tolerance vs muscle stiffness). They also noted that specific flexibility patterns are associated with specific sports or even positions within a given sport. In some instances, lesser degrees of flexibility may, in fact, be beneficial in force production associated with rapid muscular contraction (Gleim & McHugh 1997). It is likely that a variety of physiological and anthropomorphic factors allow a given athlete to obtain an elite level of performance in a given task. Traits that are beneficial for some sport-specific tasks may actually be detrimental to the performance of others. The function of an athlete needs to be viewed in the context of that athlete’s particular sport and the entire motion chain that allows for the performance of specific skills. Injury and rehabilitation must encompass this view, and deficits in flexibility (or the presence of excessive laxity) need to be assessed in the performance of a given task.

Given the data available on muscular balance, injury and soft tissue physiology, restoration of balanced strength and flexibility along the kinetic chain emerges as an important issue for rehabilitation and injury prevention. The overload process resulting in soft tissue injury and clinical symptoms is multifactorial. Aspects of technique and training clearly play a role, as do aspects of the functional kinetic chain affecting motion across a joint (Chandler & Kibler 1993). Addressing these issues in a comprehensive fashion allows for appropriate interventions to optimize athletic performance. For a detailed description of rehabilitation techniques utilizing this framework, the reader is referred to Chapter 14 in this text that is devoted to this very topic.

Conclusions

Injury and immobilization have wide-ranging physiological effects on the athlete. A comprehensive understanding of these issues is crucial in the management of an injured athlete. The athlete should be assessed not only for the acutely injured tissue but also for underlying biomechanical problems along the kinetic chain and subclinical adaptations. Athletic injuries need to be seen in the context of the entire athlete, including an accurate diagnosis of acute tissue injuries and an understanding of the individual’s clinical symptoms. The athlete should be evaluated for
other potentially overloaded structures, functional deficits such as strength imbalance, and for alterations in motion or technique that arise from or contribute to injury. Using this type of approach with a detailed knowledge of the issues discussed in other chapters of this text, clinicians can formulate a comprehensive rehabilitation strategy for the injured athlete that can restore function and provide long-term benefits in health and performance.

References


Chapter 8

Psychological Factors in Sports Injury Rehabilitation

BRITTON W. BREWER AND ALLEN E. CORNELIUS

Introduction

Rehabilitation of sports injuries involves more than repairing the physical injury and regaining previous levels of physical performance. Optimizing injury rehabilitation also includes understanding the psychological impact of the injury on the athlete and how psychological factors may interact with the rehabilitation process. The purpose of this chapter is to review the research examining psychological factors related to sports injury rehabilitation.

Major advances have been made in recent years in our understanding of the range of psychological factors associated with sports injury rehabilitation. In the early 1990s, Williams and Roepke et al. (1993) and Brewer (1994) made specific suggestions for improving the quality of psychological research on sports injury rehabilitation. Fortunately, many of these recommendations have been heeded. Specifically, responding to the recommendations made by Williams and Roepke, researchers have: (i) identified cognitive and emotional responses to injury that are characteristic of athlete populations (e.g. Leddy et al. 1994; Quinn & Fallon 1999); (ii) investigated the effects of psychological interventions on sports injury rehabilitation (e.g. Ross & Berger 1996; Theodorakis et al. 1996, 1997a, 1997b; Cupal & Brewer 2001); and (iii) initiated the education of sports medicine practitioners on psychological aspects of injury rehabilitation (e.g. Ford & Gordon 1997, 1998; Gordon et al. 1998).

Following the recommendations of Brewer (1994), researchers have: (i) advanced theory (e.g. Evans & Hardy 1995; Johnston & Carroll 1998a; Wiese-Bjornstal et al. 1998; Brewer et al. 2002); (ii) included both psychological and physical variables in their analyses (e.g. LaMott 1994; Durso-Cupal 1996; Ross & Berger 1996; Theodorakis et al. 1996; Theodorakis et al. 1997a; Theodorakis et al. 1997b; Niedfeldt 1998; Morrey et al. 1999; Brewer et al. 2000c); (iii) assessed the prevalence of clinical levels of psychological distress (Leddy et al. 1994; Brewer & Petrie 1995; Brewer et al. 1995a, 1995b; Perna et al. 1998); (iv) implemented prospective, longitudinal research designs (e.g. Smith et al. 1993; LaMott 1994; Leddy et al. 1994; Ross & Berger 1996; Morrey 1997; Petrie et al. 1997b; Udry 1997a; Perna et al. 1998; Roh et al. 1998; Morrey et al. 1999; Brewer et al. 2000c); (v) conducted qualitative investigations (e.g. Shelley 1994; Gould et al. 1997a, 1997b; Udry et al. 1997a, 1997b; Johnston & Carroll 1998a, 1998b; Bianco et al. 1999b); (vi) used control groups of athletes without injuries (e.g. LaMott 1994; Brewer & Petrie 1995; Petrie et al. 1997a; Perna et al. 1999); and (vii) examined groups of athletes that are homogeneous with respect to injury type, severity and prognosis (e.g. LaMott 1994; Ross & Berger 1996; Theodorakis et al. 1996, 1997a, 1997b; Udry 1997a; Morrey et al. 1999; Brewer et al. 2000a, 2000b, 2000c; Cupal & Brewer 2001); and (viii) used experimental research designs (e.g. Ross & Berger 1996; Theodorakis et al. 1996, 1997a, 1997b; Cupal & Brewer 2001).
To provide some structure for the voluminous amount of research that has been conducted on psychological factors in sports injury rehabilitation, two models are presented that describe the hypothesized relationships among psychological and other key variables in the sports injury rehabilitation process. The first model, the integrated model developed by Wiese-Bjornstal et al. (1998), describes many of the psychological factors related to athletes’ reactions to injuries. The second model is a biopsychosocial model developed by Brewer et al. (2002) that places sports injury rehabilitation in a broad contextual framework. Following descriptions of these models, research examining the relationships predicted by these models is reviewed.

Models of psychological response to sports injury and sports injury rehabilitation

Cognitive appraisal models

Cognitive appraisal models view an injury as a stressor or a stimulus, and the response of the individual is dependent upon a variety of factors that influence the interpretation of this stimulus. Several cognitive appraisal models have been developed to explain athletes’ reactions to sports injuries (e.g. Gordon 1986; Weiss & Troxel 1986; Grove 1993; Wiese-Bjornstal et al. 1998), the most comprehensive of which is the integrated model proposed by Wiese-Bjornstal et al. (1998) (Fig. 8.1). This model proposes that many preinjury and postinjury factors are related to how an individual reacts to a sports injury. Preinjury factors are personality, history of stressors, coping resources and interventions. Postinjury factors include personal factors (e.g. type and severity of the injury, general health status, demographic variables) and situational factors (e.g. sport played, social support system, accessibility to rehabilitation). These factors combine to determine the cognitive appraisal of the injury, which in turn affects the emotional and behavioural responses to injury, and, ultimately, the rehabilitation outcome. For example, if the cognitive response to an injury is one of doubt and thoughts of negative outcomes, a negative emotional response is likely to result. If the appraisal of the injury is more positive, with thoughts of full recovery and confidence in rehabilitation, the model predicts that a positive emotional reaction will be more likely.

There has been considerable empirical support for cognitive appraisal models in general, and the integrated model in particular. Sports injury has been identified as a significant source of stress in several studies (Brewer & Petrie 1995; Gould et al. 1997a; Bianco et al. 1999; Ford & Gordon 1999; Heniff et al. 1999), and numerous personal and situational factors have been associated with psychological responses to sports injury (Brewer 1994, 1998, 1999a). Although some of the more complex mediational relationships predicted by the integrated model have not yet found empirical support (Daly et al. 1995; Brewer et al. 2000c), there is considerable research supporting the usefulness of this model for understanding the diverse array of reactions to sports injuries. The integrated model (Wiese-Bjornstal et al. 1998) also has the benefit of providing theoretical guidance for interventions to improve rehabilitation outcomes, as many of the personal and situational factors are subject to modification (e.g. improving social support systems).

Biopsychosocial model

The integrated model describes the complex relationships of psychological, situational and cognitive variables to emotional and behavioural responses to sports injury, but it does not take into account the breadth of factors that can be related to sports injury rehabilitation processes and outcomes. Because of the multiplicity of factors that potentially interact during sports injury rehabilitation, a model that is comprehensive, yet conceptually grounded, is required. Borrowing from other health outcome research (Cohen & Rodriguez 1995; Matthews et al. 1997) and existing models of sports injury rehabilitation
Fig. 8.1 Integrated model of psychological response to the sports injury and rehabilitation process. (From Wiese-Bjornstal et al. 1998, with permission. © 1998 by the Association for the Advancement of Applied Sport Psychology.)
A model incorporating biological, social, medical and psychological factors has been developed (Brewer et al. 2002). This biopsychosocial model, as depicted in Fig. 8.2, has seven interacting components: injury characteristics, sociodemographic factors, biological factors, psychological factors, social/contextual factors, intermediate biopsychological outcomes and sports injury rehabilitation outcomes.

The biopsychosocial model proposes that the sports injury rehabilitation process begins with the occurrence of an injury. The specific characteristics of the injury (e.g. type, course, severity, location and history of injury) and sociodemographic factors (e.g. age, gender, race/ethnicity, socioeconomic status) are depicted as influencing biological factors (e.g. nutrition, sleep), psychological factors (e.g. personality, cognition) and social/contextual factors (e.g. social networks, life stress). Psychological factors are the focal point of this model, having reciprocal relationships with both biological and social/contextual factors. Biological, psychological and social/contextual factors are thought to influence intermediate biopsychological outcomes (e.g. strength, rate of recovery), with psychological factors also being reciprocally affected by these outcomes.
The final component of the model, sports injury rehabilitation outcomes, consists of functional performance, quality of life, treatment satisfaction and readiness to return to sport. This phase of the model is reciprocally related to both the intermediate biopsychological outcomes and the psychological factors.

The biopsychosocial model incorporates the various factors that are potentially related to sports injury rehabilitation outcomes, and offers a structure for examining the relationships of these factors in a contextually complete model. Given that the model incorporates findings and concepts from already established models of injury rehabilitation (Leadbetter 1994; Flint 1998; Wiese-Bjornstal et al. 1998), there is already support for many of the hypothesized relationships. However, the broad scope of the model will enable researchers to investigate a wider array of potential influences on the sports rehabilitation process.

Given the centrality of the psychological factors in the integrated model (Wiese-Bjornstal et al. 1998) and the biopsychosocial model (Brewer et al. 2002), and the focus of this chapter, the following sections review the research that examines the relationships between the psychological factors depicted in these models and the other constructs in the models. Specifically, the next section reviews the research findings related to psychological reactions to injury, and the subsequent section examines the literature concerning psychological factors associated with sports injury rehabilitation outcomes.

Psychological responses to sports injury

The subject of psychological disturbance arising from sports injury has been studied extensively since Little (1969) first identified neurotic symptoms in injured athletes. The following sections examine research concerning emotional and behavioural responses to sports injury that has accrued since Little’s seminal work. For the purposes of this discussion, sports injuries are defined as physical damage that occurs as the result of an event or process associated with sports participation and disrupts subsequent sports participation. The duration of such disruption of involvement in sports may vary extensively depending on the nature, course and severity of the injury. Emotional responses refer to affective (i.e. feeling) states that are experienced subsequent to injury onset, whereas behavioural responses refer to overt actions that are manifested following the occurrence of injury.

Emotional responses to sports injury

Emotional responses to sports injury have been a widely researched topic within the sports psychology literature. A popular approach has been the application of stage theories that propose that athletes experience an injury as a type of loss. The injury is thought to be perceived by the athlete as a loss of an aspect of the self, and therefore considered to elicit a reaction similar to that exhibited by an individual who has experienced a serious psychological loss (Peretz 1970; Rotella & Heyman 1986). Stage models of grief and loss have proposed that individuals experience a sequence of emotional responses to the loss, leading to eventual adaptation. In perhaps the best known description of these stages, Kubler-Ross 1969) proposed that in response to a major loss, individuals commonly progress through five stages: denial, anger, bargaining, depression and acceptance. Researchers have found that athletes experience many of these same emotions following an injury (Rotella 1985; Astle 1986; Lynch 1988; Silva & Hardy 1991).

Although the five-stage model developed by Kubler-Ross (1969) has been widely discussed by sport psychologists, research investigating athletes’ reactions to injury has not fully supported this model. Kubler-Ross’s model was derived from her work with terminally ill patients, and the experience of being terminally ill may not be similar enough to having a sports injury to be generalizable (Smith et al. 1990b). There is, however, support for emotional responses to injury being somewhat similar to grief reactions
psychological factors (Shelley 1994; Macchi & Crossman 1996). In the initial time period following an injury, negative emotions such as depression, frustration, confusion, anger and fear have been documented in several qualitative studies (Gordon & Lindgren 1990; Shelley & Carroll 1996; Shelley & Sherman 1996; Udry et al. 1997a; Johnston & Carroll 1998a; Sparkes 1998; Bianco et al. 1999b). During the middle phase of rehabilitation, depression and frustration are commonly reported emotions, with the source of these emotions shifting from concerns about the injury to rehabilitation-related issues. As rehabilitation of the injury nears completion, depression and frustration remain prevalent and a fear of reinjury emerges (Johnston & Carroll 1998a; Bianco et al. 1999b).


A deviation from this pattern, which has been documented in two studies (LaMott 1994; Morrey et al. 1999), is a slight increase in negative emotions and a slight decrease in positive emotions as athletes near the end of rehabilitation following reconstructive knee surgery, possibly reflecting athletes’ apprehension about returning to sports activity and fear of reinjury (Quinn 1996; Johnston & Carroll 1998a; Bianco et al. 1999b). Stage models ignore several important factors related to how athletes respond to an injury, such as their idiosyncratic perceptions of the injury, and situational circumstances such as the severity of the injury and sources of social support (Brewer 1994; Wiese-Bjornstal et al. 1998). The general lack of empirical support for stage models does not mean that they are completely without merit. Recently proposed versions of stage models have evolved from static and rigid stages of previous models to more flexible and dynamic descriptions of psychological responses to sports injury, thereby allowing for more individual variation in how the stages are experienced (Evans & Hardy 1995, 1999). Such modification muddles the simplicity of stage models, but...
reflects more accurately the dynamic nature of the emotional experience of athletes who have been injured. Stage models have provided impetus for the development of cognitive appraisal models of responses to injury, such as the integrated model (Wiese-Bjornstal et al. 1998), which identifies factors that mediate and moderate emotional responses to sports injury.

**Mediating and moderating factors**

The integrated model (Wiese-Bjornstal et al. 1998) posits cognitive, personal and situational factors as influences on emotional and behavioural responses to sports injury, with cognitive appraisals mediating the relationships between personal and situational factors and emotional and behavioural responses. However, most of the research in this area has not addressed this mediational effect, but has focused instead on direct associations between various personal and situational factors and emotional and behavioural responses to injury. Therefore, only indirect evidence of the hypothesized mediational role of cognitive appraisals has been obtained.

Personal factors that have been identified as being positively related to postinjury emotional disturbance include self-identification with the athlete role (Brewer 1993; Shelley & Carroll 1996; Sparkes 1998), competitive trait anxiety (Petrie et al. 1997), investment in playing professional sports (Kleiber & Brock 1992), level of sports involvement (Meyers et al. 1991), pessimistic explanatory style (Grove et al. 1990) and previous injury experience (Bianco et al. 1999b). Athletes with injuries have demonstrated negative relationships between emotional distress and age (Smith et al. 1990a; Brewer et al. 1995a) and hardiness (Grove et al. 1990; Miller 1998). Thus, the individuals most likely to encounter difficulty adjusting emotionally to injury are those who are young, least hardy, most strongly identified with the athlete role, most dispositionally anxious, most invested in having a career as a professional athlete, most experienced in the rigors of sports injury rehabilitation, and most pessimistic. The relationship of emotional distress to age is potentially more complex than a simple linear relationship, as Meyers et al. (1991) obtained a curvilinear relationship between age and emotional disturbance for participants recovering from knee surgery. Participants of intermediate age (20–39 years) reported greater levels of emotional disturbance than younger (10–19 years) and older (40–49 years) participants. Injury-related characteristics and aspects of the social and physical environments that can change over time are referred to as situational factors. As shown in Table 8.1, numerous situational factors have been correlated with emotional distress in athletes.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Direction</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current injury status</td>
<td>+</td>
<td>Alzate et al. 1998; Brewer et al. 1995a; Quinn 1996</td>
</tr>
<tr>
<td>Injury severity</td>
<td>+</td>
<td>Alzate et al. 1998; Pargman &amp; Lunt 1989; Perna 1992; Smith et al. 1990a, 1993</td>
</tr>
<tr>
<td>Impairment of daily activities</td>
<td>+</td>
<td>Crossman &amp; Jamieson 1985</td>
</tr>
<tr>
<td>Life stress</td>
<td>+</td>
<td>Brewer 1993; Petrie et al. 1997b; Quinn 1996</td>
</tr>
<tr>
<td>Recovery progress</td>
<td>+</td>
<td>Quinn 1996</td>
</tr>
<tr>
<td>Impairment of sport activities</td>
<td>–</td>
<td>Brewer et al. 1995a</td>
</tr>
<tr>
<td>Level of sport involvement</td>
<td>+/-</td>
<td>Crossman et al. 1995; Meyers et al. 1991</td>
</tr>
<tr>
<td>Medical prognosis</td>
<td>–</td>
<td>Gordin et al. 1988</td>
</tr>
<tr>
<td>Recovery progress</td>
<td>–</td>
<td>McDonald &amp; Hardy 1990; Quinn 1996; Smith et al. 1988</td>
</tr>
<tr>
<td>Social support for rehabilitation</td>
<td>–</td>
<td>Brewer et al. 1995a</td>
</tr>
<tr>
<td>Social support satisfaction</td>
<td>–</td>
<td>Green &amp; Weinberg 1998</td>
</tr>
</tbody>
</table>

Direction refers to the positive (+) and negative (−) correlations with postinjury emotional disturbance.
with injury. In general, research has indicated that athletes are likely to experience greater postinjury emotional disturbance when they perceive their injuries as serious, view themselves as having made little rehabilitative progress, and consider themselves as weakly supported in their rehabilitative pursuits.

According to the integrated model (Wiese-Bjornstal et al. 1998), it is cognitive appraisals made concerning the injury that have the most direct effect on emotional and behavioural responses. Numerous types of cognitive appraisals have been correlated with emotional distress in research studies. Greater emotional distress following injury has been associated with the tendencies to interpret pain in a catastrophic manner (Tripp 2000) and to attribute the cause of sports injury to factors residing within oneself (Tedder & Biddle 1998) and pertaining to all areas of one’s life (Brewer 1991). Lower levels of postinjury emotional distress have been found for individuals who: (i) think that they will be able to cope with their injuries (Daly et al. 1995); (ii) can readily imagine themselves functioning favourably following injury (Fisk & King 1998); (iii) have high general and physical self-esteem (Brewer 1993; Quinn 1996); and (iv) are confident in themselves and their ability to adhere to the rehabilitation protocol, recover fully from injury and succeed in sport (Quinn 1996). Also, in contrast to the findings of Tedder and Biddle, athletes who attribute the cause of their injuries to factors residing within themselves and likely to occur have reported lower levels of postinjury emotional disturbance (Brewer 1999b). Thus, although it is unclear whether taking responsibility for one’s injury is adaptive in terms of emotional adjustment, there is little doubt that cognitions indicating confidence in oneself, one’s body and one’s recovery are associated with more favourable emotional states following sports injury.

Behavioural responses to sports injury

A sports injury can elicit a number of behavioural responses in addition to the emotional responses discussed in the previous section. The two most relevant behaviours, and the most researched, are coping behaviours and adherence to sports injury rehabilitation.

Coping behaviours

In an effort to deal effectively with injuries and the related issues that may arise during rehabilitation, athletes may engage in certain behaviours to help them cope with the situation. Specific behavioural strategies that have been identified in athletes with injuries include an aggressive rehabilitation approach, avoiding others, building strength, distracting oneself (e.g. keeping busy, seeking a change of scenery), ‘driving through’ (e.g. doing things normally, learning about their injuries, resting when tired, working hard to achieve rehabilitation goals), seeking out and using social support networks, trying alternative treatments, and working or training at their own pace (Gould et al. 1997b; Bianco et al. 1999a).

Other studies have examined the use of specific coping strategies, and have found that the most strongly endorsed coping behaviours assessed by the COPE inventory (Carver et al. 1989; Scheier et al. 1994) were active coping, which involves initiating behaviours to deal directly with a stressor or its effects, and instrumental social support, which pertains to seeking help or information (Grove & Bahnsen 1997; Udry 1997a; Quinn & Fallon 1999).

Adherence to sports injury rehabilitation

A behavioural response to sports injury that is of central focus to the rehabilitation of the injury is adherence to the recommended rehabilitation protocol, which can involve a number of different behaviours, including: (i) performing clinic-based activities, such as doing exercises designed to increase strength, flexibility and endurance; (ii) modifying physical activity, such as resting and limiting activity; (iii) taking medications; and (iv) completing home-based activities, such as cryotherapy and home rehabilitation exercises.
(Brewer 1998, 1999a). This wide range of behaviours requires a correspondingly wide range of assessment techniques for measuring adherence. The most frequently used methods of assessing adherence are documenting patient attendance at clinic-based rehabilitation sessions, recording practitioner ratings of adherence during rehabilitation sessions, and obtaining patient self-reports of home exercise completion (Brewer 1999a). Complicating this diversity of adherence measures is a lack of consistent operationalization of adherence behaviour, with some research reporting a percentage of adherent versus non-adherent individuals, and other research comparing actual adherence behaviour to that recommended by the practitioner.

Recognizing this diversity of adherence measures and conflicting operationalization of the adherence construct, estimates of adherence to sports injury rehabilitation have ranged from 40 to 91% (Brewer 1998, 1999a). Adherence rates tend to be higher for continuous measures of adherence, such as attendance at rehabilitation sessions (e.g. Almekinders & Almekinders 1994; Daly et al. 1995; Laubach et al. 1996) or amount of time spent on home rehabilitation activities (Penpraze & Mutrie 1999), than for more discrete measures of adherence that categorize individuals based on their level of adherence (e.g. Taylor & May 1996). Consistent with a cognitive appraisal approach and the integrated model (Wiese-Bjornstal et al. 1998), researchers have postulated that adherence rates are related to personal and situational factors, as well as cognitive and emotional responses. However, as with emotional responses to sports injury, there is little direct support for the hypothesized mediational role of cognitive appraisals in the relations between personal and situational factors and adherence to sports injury rehabilitation. But researchers have identified many personal, situational and cognitive factors that are associated with sports injury rehabilitation adherence.

Personal factors that have been identified as having a positive relationship with sports injury rehabilitation adherence include: (i) an internal health locus of control (Murphy et al. 1999; Brewer et al. 2000b), which refers to the extent to which a person believes that health outcomes are under their own control; (ii) pain tolerance (Fisher et al. 1988; Byerly et al. 1994; Fields et al. 1995); (iii) self-motivation (Noyes et al. 1983; Fisher et al. 1988; Duda et al. 1989; Brewer et al. 1994a; Fields et al. 1995; Culpepper et al. 1996; Brewer et al. 2000c); (iv) task involvement (Duda et al. 1989), which is the degree to which a person is motivated to improve against their own personal standards; and (v) tough mindedness (Wittig & Schurr 1994). Personal factors that have been shown to be negatively associated with adherence to sports injury rehabilitation are: (i) a chance health locus of control (Brewer et al. 2000b), which refers to the extent to which a person believes that health outcomes are influenced by chance or luck; (ii) ego involvement (Duda et al. 1989), which is the degree to which a person is motivated by comparisons with other individuals; and (iii) trait anxiety (Eichenhofer et al. 1986). Thus, in terms of dispositional characteristics, the athletes most likely to adhere to their rehabilitation programmes are those who are self-motivated, strong-willed and tolerant of discomfort.

As shown in Table 8.2, a diverse array of situational factors has been found to correlate with adherence to sports injury rehabilitation programmes. In general, higher levels of adherence have been associated with: (i) a supportive clinical environment; and (ii) a rehabilitation programme that is convenient, valued and perceived as efficacious.

The integrated model (Wiese-Bjornstal et al. 1998) predicts that, as with personal and situational factors, cognitive variables should be related to behaviours such as adherence to rehabilitation. Research has shown that individuals who have high adherence rates also report a high ability to cope with their injuries (Daly et al. 1995), have high rehabilitation self-efficacy (Taylor & May 1996), perceive little threat to their self-esteem (Lampton et al. 1993), attribute their recovery to stable and personally controllable factors (Laubach et al. 1996), set rehabilitation goals, use imagery and use positive self-talk (Scherzer et al. 1999).
Thus, athletes who are confident in their ability to meet the demands of rehabilitation, accept responsibility for their rehabilitation, and allocate mental effort to their rehabilitation tend to exhibit the highest levels of adherence.

One of the few cognitive processes that has received any experimental attention with regard to sports injury rehabilitation adherence is goal setting. Penpraze and Mutrie (1999) assigned athletes with injuries to either a group that was assigned specific rehabilitation goals or a group that received non-specific rehabilitation goals. Athletes in the specific goals group had a greater understanding of, and adherence to, their rehabilitation protocols than athletes in the non-specific goal group. These experimental findings extended earlier qualitative findings that task-oriented goal setting was related to a greater perception of rehabilitation adherence in athletes with injuries (Gilbourne & Taylor 1995, 1998; Gilbourne et al. 1996).

The integrated model also predicts that emotional responses and other behavioural responses will be associated with adherence to sports injury rehabilitation, although only a few studies have examined such relationships. Mood disturbance has been negatively related to adherence in three (Daly et al. 1995; Brickner 1997; Alzate et al. 1998), but a fourth study found no association between adherence levels and psychological disturbance (Brewer et al. 2000c). The only behavioural response that has been investigated in association with sports injury rehabilitation adherence has been the use of instrumental coping behaviours, which involve asking for additional information about the injury or the rehabilitation programme (Udry 1997a). Instrumental coping behaviours were positively related to adherence levels for individuals who were undergoing rehabilitation following reconstructive knee surgery designed to facilitate a return to sports participation.

### Psychological factors and sports injury rehabilitation outcomes

In addition to psychological responses to sports injury reviewed in the preceding section, the psychological correlates of sports injury rehabilitation outcomes are also of vital interest. Both

<table>
<thead>
<tr>
<th>Variable</th>
<th>Direction</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic class status</td>
<td>+</td>
<td>Culpepper et al. 1996; Shank 1988</td>
</tr>
<tr>
<td>Academic performance level</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Belief in the efficacy of the treatment</td>
<td>+</td>
<td>Duda et al. 1989; Noyes et al. 1983; Taylor &amp; May 1996</td>
</tr>
<tr>
<td>Comfort of the clinical environment</td>
<td>+</td>
<td>Brewer et al. 1994a; Fields et al. 1995; Fisher et al. 1988</td>
</tr>
<tr>
<td>Convenience of rehabilitation scheduling</td>
<td>+</td>
<td>Fields et al. 1995; Fisher et al. 1988</td>
</tr>
<tr>
<td>Degree of career goal definition</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Importance or value of rehabilitation</td>
<td>+</td>
<td>Taylor &amp; May 1996</td>
</tr>
<tr>
<td>Injury duration</td>
<td>+</td>
<td>Culpepper et al. 1996</td>
</tr>
<tr>
<td>Perceived academic load</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Perceived sports participation time</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Perceived time availability for rehabilitation</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Perceived exertion during rehabilitation</td>
<td>+</td>
<td>Brewer et al. 1994a; Fisher et al. 1988</td>
</tr>
<tr>
<td>Perceived susceptibility to further</td>
<td>+</td>
<td>Taylor &amp; May 1996</td>
</tr>
<tr>
<td>complications without rehabilitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcollegiate sports participation plans</td>
<td>+</td>
<td>Shank 1988</td>
</tr>
<tr>
<td>Rehabilitation practitioner expectancy</td>
<td>+</td>
<td>Taylor &amp; May 1995</td>
</tr>
<tr>
<td>Social support for rehabilitation</td>
<td>+</td>
<td>Byerly et al. 1994; Duda et al. 1989; Finnie 1999; Fisher et al. 1988</td>
</tr>
</tbody>
</table>

Direction refers to the positive (+) correlations with postinjury emotional disturbance.
the integrated model (Wiese-Bjornstal et al. 1998) and the biopsychosocial model (Brewer et al. 2002) specify relationships between psychological variables and sports injury rehabilitation outcomes. These models predict that the same clusters of personal, situational, cognitive, emotional and behavioural variables associated with psychological responses to sports injury will also be related to sports injury rehabilitation outcomes. Research that has investigated relationships between psychological factors and sports injury rehabilitation outcomes is discussed in the next sections. For the purposes of this discussion, sports injury rehabilitation outcomes refer not only to the variables listed in the category of the same name in the biopsychosocial model shown in Fig. 8.2, but also to the variables listed in the biopsychosocial model (Brewer et al. 2002) as ‘intermediate biopsychological outcomes’, as these variables (e.g. endurance, joint laxity, pain, range of motion, recovery rate, strength) are often used as indices of rehabilitation outcome in research investigations.

PERSONAL FACTORS

Personal factors related to sports injury rehabilitation outcomes have been a topic of interest since Wise et al. (1979) discovered that two personality variables (i.e. hysteria and hypochondriasis) were inversely related to recovery following knee surgery. Subsequent studies have shown that being optimistic (LaMott 1994), male (Johnson 1996, 1997) and strongly identified with the athlete role (Brewer et al. 2000c) have all been positively related to rehabilitation outcomes following sports injuries. Research has not yet uncovered the mechanisms by which these personal factors may influence outcome, or what other of the myriad personal factors may be related to sports injury rehabilitation outcomes.

SITUATIONAL FACTORS

The situational factor that has received the most amount of research attention with respect to sports injury rehabilitation outcomes has been social support, which refers to the quantity, quality and type of interactions that athletes have with other people (Udry 1996). Social support is considered to be multidimensional, with the structure proposed by Richman et al. (1993) the most widely used in sports injury rehabilitation research (Ford & Gordon 1993; Izzo 1994; LaMott 1994; Bianco & Orlick 1996; Quinn 1996; Ford 1998; Johnston & Carroll 1998b). The Richman structure categorizes social support as one of eight types: listening support, emotional support, emotional challenge, task appreciation, task challenge, reality confirmation, material assistance and personal assistance. These different types of social support may be provided by different members of an athlete’s social network, including coaches, friends, relatives, team mates, sports administrators and medical personnel (Lewis & LaMott 1992; Izzo 1994; Bianco & Orlick 1996; Macchi & Crossman 1996; Peterson 1997; Udry 1997b; Udry et al. 1997b; Ford 1998; Johnston & Carroll 1998b; Bianco et al. 1999a; Udry & Singleton 1999). The needs for particular types of social support and the ability of an athlete’s social support network to provide certain types of support may vary during the rehabilitation process (Ford 1998; LaMott 1994; Quinn 1996; Udry 1997a; Johnston & Carroll 1998b).

Given the complexity of social support networks and the changing social support needs of injured athletes over the course of rehabilitation, it is not surprising that research in this area has produced conflicting results. For example, social support has been positively related (Tuffey 1991), not related (Brewer et al. 2000c) and negatively related (Quinn & Fallon 2000) to rehabilitation outcome. These inconsistencies are probably due to the way in which social support has been differentially operationalized across different studies. The importance of considering different aspects of social support and their potential different relationships to rehabilitation outcome is highlighted in a study of skiers (Gould et al. 1997a). Skiers who experienced successful injury rehabilitation were less likely to perceive a lack of attention/empathy from others and less likely to encounter negative social relationship, yet
were more likely to indicate feeling socially isolated than skiers who experienced unsuccessful injury rehabilitation. Thus, when examining the relationships between rehabilitation outcome and social support, it is imperative that the specific type of social support is considered. Further research is needed to better understand the roles of the different types of social support in the injury rehabilitation process.

**COGNITIVE FACTORS**

As indicated in Table 8.3, research has identified many cognitive factors that are associated with sports injury rehabilitation outcomes. Overall, the research in this area suggests that positive cognitions and the use of psychological skills will enhance the rehabilitation process. Some of the findings, however, such as those for denial and emotion-focused coping/emotional focus (Quinn 1996; Grove & Bahnsen 1997; Quinn & Fallon 2000), are equivocal. These inconsistencies, along with the retrospective and correlational nature of the majority of studies on cognitive factors associated with sports injury rehabilitation outcomes, indicate that more research (especially prospective and experimental research) is needed in this area to better understand the role of cognition in influencing rehabilitation outcomes.

**EMOTIONAL FACTORS**

Only a few studies have examined relationships between emotional variables and sport rehabilitation outcomes. Positive relationships to have been found between rehabilitation outcomes and general well-being (Johnson 1996, 1997) and vigour (Quinn 1996), whereas negative associations have been documented for anger (LaMott 1994; Alzate et al. 1998), anxiety (Johnson 1996, 1997), fear, frustration, relief (LaMott 1994), mood disturbance, depression (Alzate et al. 1998; Tripp 1997), and...
2000), fatigue, tension (Alzate et al. 1998) and psychological distress (Brewer et al. 2000c). Although all of the findings regarding the relationship between emotions and sports injury rehabilitation outcomes are purely correlational in nature, most of the studies cited (i.e. Johnson 1996, 1997; Alzate et al. 1998; Brewer et al. 2000c; Quinn & Fallon 2000) were prospective in that emotions were measured at one point in time (e.g. prior to surgery, at the beginning of rehabilitation) and sports injury rehabilitation outcomes were measured at a later point in time (e.g. at the end of rehabilitation). Consequently, there is evidence that for reasons not currently understood, positive emotions may often precede favourable rehabilitation outcomes.

BEHAVIOURAL FACTORS

The behavioural factor examined most frequently in reference to sports injury rehabilitation outcomes has been adherence to rehabilitation. One would assume that greater adherence to rehabilitation would be associated with better outcome, but this has not always been the case. The relationship between adherence to rehabilitation and rehabilitation outcome has been found to be positive (Meani et al. 1986; Derscheid & Feiring 1987; Hawkins 1989; Satterfield et al. 1990; Tuffey 1991; Treacy et al. 1997; Alzate et al. 1998; Brewer et al. 2000c; Quinn & Fallon 2000), non-significant (Noyes et al. 1983; Brewer et al. 2000c; Quinn & Fallon 2000) and, surprisingly, negative (Shelbourne & Wilckens 1990; Quinn & Fallon 2000). These discrepant findings are probably due to a variety of factors, including the nature of the injury studied, the specifics of the rehabilitation protocol, the phase of rehabilitation that was the focus of study, and the particular measures of adherence and outcome (Brewer 1999a).

Only a few behaviours other than adherence to rehabilitation have received any empirical investigation. Better sports injury rehabilitation outcomes have been found to be associated with higher levels of active coping (Quinn & Fallon 1999), lower levels of physical activity (Gould et al. 1997b) and higher levels of seeking social support (Johnson 1996, 1997; Gould et al. 1997b).

PSYCHOLOGICAL INTERVENTIONS TO ENHANCE REHABILITATION OUTCOMES

Support for the influence of a wide variety of psychological factors on sports injury rehabilitation outcomes can be inferred from the results of experimental studies in which psychological interventions have been applied to athletes with injuries. As shown in Table 8.4, a number of

<table>
<thead>
<tr>
<th>Reference</th>
<th>Intervention</th>
<th>Effect(s) of intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krebs 1981</td>
<td>Biofeedback</td>
<td>Greater strength and EMG output</td>
</tr>
<tr>
<td>Draper 1990</td>
<td>Biofeedback</td>
<td>Greater strength and ROM</td>
</tr>
<tr>
<td>Draper &amp; Ballard 1991</td>
<td>Biofeedback</td>
<td>Greater strength</td>
</tr>
<tr>
<td>Levitt et al. 1995</td>
<td>Biofeedback</td>
<td>Greater extensor torque and quadriceps fibre recruitment</td>
</tr>
<tr>
<td>Theodorakis et al. 1996</td>
<td>Goal setting</td>
<td>Greater strength and self-efficacy</td>
</tr>
<tr>
<td>Theodorakis et al. 1997b</td>
<td>Goal setting</td>
<td>Greater physical functioning, less pain, less reinjury anxiety</td>
</tr>
<tr>
<td>Ross &amp; Berger 1996</td>
<td>Stress inoculation training</td>
<td>Greater strength</td>
</tr>
<tr>
<td>Theodorakis et al. 1997b</td>
<td>Self-talk</td>
<td>Greater positive mood and readiness for competition</td>
</tr>
<tr>
<td>Johnson 2000</td>
<td>Multimodal</td>
<td>Greater strength, less pain, less reinjury anxiety</td>
</tr>
<tr>
<td>Cupal &amp; Brewer 2001</td>
<td>Relaxation/guided imagery</td>
<td>Greater strength</td>
</tr>
</tbody>
</table>
Interventions have been investigated experimentally for their effects on physical and psychological rehabilitation outcomes, including biofeedback (Krebs 1981; Draper 1990; Draper & Ballard 1991; Levitt et al. 1995), goal setting (Theodorakis et al. 1996, 1997b), imagery/relaxation (Cupal & Brewer 2001), self-talk (Theodorakis et al. 1997a), stress inoculation training (Ross & Berger 1996) and a multimodal intervention consisting of goal setting, imagery, relaxation and stress management (Johnson 2000). Case study data have also supported the efficacy of interventions such as counselling, goal setting, hypnosis, positive self-talk, relaxation and systematic desensitization for positively effecting rehabilitation outcome variables such as confidence, motivation, perception of pain, physical recovery, psychological adjustment, reinjury anxiety and range of motion (Rotella & Campbell 1983; Nicol 1993; Sthalekar 1993; Potter 1995; Brewer & Helledy 1998; Hartman & Finch 1999; Evans et al. 2000; Jevon & O’Donovan 2000).

Without exception, the psychological interventions that have been documented in the scientific literature as having been used successfully in sports injury rehabilitation are cognitive-behavioural in nature and involve athletes learning new skills or behaviours to cope more effectively with the rehabilitation process, both physically and psychologically. Biofeedback, for example, involves furnishing patients with physiological information (e.g. electromyographic activity in the quadriceps muscle group) and is thought to produce therapeutic gains by enhancing motivation for rehabilitation activities and enhancing proprioceptive information processing (Levitt et al. 1995). Similarly, goal setting, which involves generating (short- and long-term) personal standards of achievement in rehabilitation activities, is posited to provide direction to the athlete’s rehabilitation efforts, enhance persistence in rehabilitation and facilitate the development of new rehabilitation strategies (Locke & Latham 1990). Although they are conceptually distinct interventions, relaxation and imagery are often used in conjunction to treat athletes with injuries. Relaxation protocols, which are typically aimed at calming the muscles and the mind, are frequently implemented just prior to the use of imagery procedures, which generally feature mental rehearsal of motivational, healing and performance aspects of rehabilitation. Combining relaxation and imagery techniques is thought to affect rehabilitation outcomes by enhancing rehabilitation motivation and boosting physiological processes such as tissue regeneration/repair and immune/inflammatory responses (Cupal & Brewer 2001). Self-talk (or cognitive restructuring) interventions are designed to alter athletes’ thoughts and, ultimately, feelings and behaviours, regarding their rehabilitation.

Because research on psychological interventions in sports injury rehabilitation has focused primarily on documenting the effectiveness of such interventions, the processes by which these interventions exert their effect are not well understood. Referring to the biopsychosocial model, it is postulated that psychological interventions may affect rehabilitation outcome through a variety of mechanisms, including a direct effect, indirect effects through intermediate biopsychological outcomes, and indirect effects through relationships with biological factors and social/contextual factors. The complexities of these possible relationships have yet to be adequately explored by researchers, but it is likely that psychological interventions affect outcomes through a variety of pathways.

Implications for clinical practice

The percentage of athletes who experience difficulties adjusting emotionally or behaviourally to injury furnishes evidence of the need to consider psychological factors in planning, implementing and evaluating sports injury rehabilitation protocols. Further rationale for incorporating psychological aspects into the treatment of sports injuries is provided by the abundance of psychological factors associated with sports injury rehabilitation outcomes and the demonstrated efficacy of psychological interventions in
producing desirable sports injury rehabilitation outcomes. Although the value of taking psychological factors into account in sports injury rehabilitation is recognized, the ways in which psychology should be included in the rehabilitation process and by whom is less clearly delineated. Potential opportunities for clinical application of the research on psychological factors in sports injury rehabilitation exist for both sports injury rehabilitation practitioners and sport psychology professionals.

SPORTS INJURY REHABILITATION PRACTITIONERS

A variety of medical professionals provide clinical services to athletes with injuries, including physicians, physiotherapists and athletic trainers. Depending on the frequency of their contact with athletes during the injury rehabilitation process, sports injury rehabilitation practitioners can be in a strategic position to enhance the adjustment of athletes to injury and promote adherence to the treatment protocol. Indeed, sports injury rehabilitation practitioners, in constituting an important situational factor in the integrated model (Wiese-Bjornstal et al. 1998) and social/contextual variable in the biopsychosocial model (Brewer et al. 2002), have the potential to affect the psychological state of their patients regardless of their intent to do so.

Surveys of sports injury rehabilitation practitioners (e.g. Brewer et al. 1991; Gordon et al. 1991; Ford & Gordon 1993, 1997; Larson et al. 1996; Ninedek & Kolt 2000) have indicated a general awareness of, and interest in, psychological aspects of rehabilitation. Despite this interest and awareness, sports injury practitioners may not feel comfortable or qualified to make judgements or initiate interventions of a psychological nature. Research has indicated that sports injury rehabilitation practitioners may have difficulty recognizing psychological distress among their patients (Brewer et al. 1995b; Maniar et al. 1999). With appropriate training (Ford & Gordon 1998; Gordon et al. 1998), however, rehabilitation professionals can increase their knowledge, skills and comfort in identifying the warning signs of athletes who are experiencing problems in adjusting psychologically to their injuries.

In some cases, such as when high levels of depression or anxiety are detected, referral to a mental health practitioner is warranted (Brewer et al. 1999b). In most circumstances, though, sports injury rehabilitation practitioners can facilitate a smooth navigation of the rehabilitation process by educating athletes about their injuries and likely challenges to be encountered during rehabilitation and, when appropriate, applying a simple psychological intervention such as goal setting (Gilbourne & Taylor 1995, 1998; Gilbourne et al. 1996; Theodorakis et al. 1996, 1997; Penpraze & Mutrie 1999) to motivate and focus the rehabilitation efforts of athletes. The importance of patient education and clear patient–practitioner communication is underscored by research demonstrating that athletes with injuries often misperceive their interactions with rehabilitation practitioners (Kahanov & Fairchild 1994) and misunderstand at least some portion of their prescribed rehabilitation regime (Webborn et al. 1997).

SPORT PSYCHOLOGY PROFESSIONALS

Despite the growing body of literature documenting the role of psychological factors in influencing sports injury rehabilitation processes and outcomes, only rarely are sport psychology professionals fully fledged members of the sports medicine treatment team (Cerny et al. 1992; Larson et al. 1996). The low level of involvement of sport psychology professionals in the rehabilitation of athletes with injuries may be due, in part, to: (i) the structure and restrictions of the health care systems in which athletes receive rehabilitation for their injuries; (ii) the lack of standard procedures for rehabilitation practitioners to refer athletes with injuries for counselling or psychotherapy (Larson et al. 1996); and (iii) the reluctance of athletes with injuries to participate in ‘extra’ therapeutic activity beyond their physical rehabilitation even if they find the interventions sufficiently credible (Brewer et al. 1994b).
In circumstances where a sport psychology professional is not involved in the day-to-day assessment of athletes with injuries and in the initiation of psychological interventions such as imagery and relaxation, as seems to be the norm, sports injury rehabilitation practitioners are encouraged to identify sport psychology professionals to serve as consultants, referral targets and resources on psychological aspects of rehabilitation. Some sport psychology professionals have clinical training and can work with athletes who show signs of mental disorders or severe adjustment reactions. Unfortunately, when sport psychology professionals are not integral members of the sports medicine treatment team, they are more likely to be contacted regarding difficult patients than regarding routine ways that psychology can be applied to enhance the rehabilitation of athletes with injuries. Further research documenting the role of psychological factors and the efficacy of psychological interventions in sports injury rehabilitation is needed to advance the standing of sport psychology professionals in the enterprise of sports injury rehabilitation.

**Future research directions**

The proliferation of, and progress in, research concerning psychological correlates of sports injury rehabilitation parameters since the recommendations made by Williams and Roepke (1993) and Brewer (1994) has been heartening. However, this body of research has only scratched the surface of the potential questions still in need of answers. The two models used as a framework for the review of literature in this chapter, the integrated model (Wiese-Bjornstal et al. 1998) and the biopsychosocial model (Brewer et al. 2002), have provided not only a structure for examining the literature to date, but offer many suggestions for future research. As demonstrated by the research reviewed in this chapter, many of the postulated direct pathways of both models have been supported, but many have yet to be examined. In particular, although both models predict that psychological factors will influence emotions, behaviour, cognitions and, ultimately, rehabilitation outcome, most studies have focused on highly specific psychological factors and only one of these types of variables. This has led to a hodgepodge of direct relationships, with little comparability of the psychological factors studied across the different outcome measures. Both models also propose that many psychological factors are related to injury and injury rehabilitation processes as mediating variables, and these mediating relationships have yet to be sufficiently investigated. To examine these relationships, studies will have to become more integrated, including measures of the multiple constructs of both the integrated model and the biopsychosocial model. It must also be recognized that the rehabilitation process is a dynamic one, and more frequent assessments of the psychological factors (as in time series analysis) and intrapersonal analyses are needed to examine the fluid and diverse nature of the constructs involved (Evans & Hardy 1999).

One area of research vital to ensure progress in this area of enquiry is the development and standardization of adequate measures of the specific constructs involved. Diversity in measurement methods and inconsistent operationalization of constructs has led to difficulty in making comparisons across studies. Increased standardization of instruments and definitions of the constructs are likely to lead to a more consistent and unified body of research (Evans & Hardy 1999). However, the development of adequate instruments is a difficult and time-consuming endeavour, and research on psychological aspects of sports injury rehabilitation has suffered from instruments that have not demonstrated good psychometric properties (Brewer et al. 1999a; Slattery 1999). Only with adequately designed and validated measures can the various constructs involved in the psychological aspects of sports injury be effectively examined.

A further area of enquiry on which future research should focus is the specific processes and effectiveness of psychological interventions in sports injury rehabilitation. Although a variety of interventions have been found to be related
to better sports injury rehabilitation outcomes (Cupal 1998), the specifics of the mechanisms through which these interventions operate are not well understood. Only by clearly understanding how these interventions work can the interventions be tailored effectively to the multiplicity of personal variables, situational variables and psychological variables implicated in the sports injury rehabilitation process.

It will also be important to examine ways in which psychological interventions interact with medical interventions (e.g. immobilization, medications, physical therapy, surgery, therapeutic exercise) to influence emotional and behavioural functioning during injury rehabilitation. Given that, from a biopsychosocial perspective (Brewer et al. 2002), medical interventions may directly affect both biological factors and intermediate biopsychosocial outcomes, it would not be surprising if certain psychological interventions were found to be especially effective for some injuries and in conjunction with particular medical treatments. For example, a psychological intervention designed to enhance motivation might be more appropriate for injury rehabilitation protocols that have heavy behavioural demands (e.g. extensive rehabilitation exercises) than for those with more passive features (e.g. activity restriction, immobilization). Unquestionably, the agenda for further research on psychological interventions in sports injury rehabilitation should include randomized controlled clinical trials comparing frequently advocated psychological interventions (e.g. cognitive restructuring, counseling, goal setting, imagery, relaxation) separately and in combination. Such research can help provide practitioners with more detailed information on what interventions, administered in which ways and by whom at which frequencies, are likely to be most effective for which athletes under what circumstances.

A paradoxical direction for future research is to explore the potential benefits of sports injury (Udry 1999). Some recent qualitative studies have documented that some athletes with injuries have reported experiencing higher levels of life satisfaction, performance enhancement, personal growth, psychosocial development and academic performance as a result of their injuries (Rose & Jevne 1993; Udry et al. 1997a; Ford 1998; Niedfeldt 1998; Ford & Gordon 1999). However, again, the specific personal, situational and psychological variables that are related to these benefits have not been identified, and further studies, both qualitative and quantitative, are needed to expand our knowledge of this fascinating counter-intuitive line of research.

From a theoretical standpoint, explanatory models for psychological aspects of sports injury are in their infancy. The integrated model (Wiese-Bjornstal et al. 1998) and the biopsychosocial model (Brewer et al. 2002) described in this chapter have been proposed only recently, and although supported by a growing body of research, there are still many predictions derived from these models that have not yet received empirical support. Further data gathered within the contexts of these models will probably inspire modification to the models as they currently exist. Additionally, there is evidence indicating that psychological factors are related to the occurrence of injury (Meeuwisse & Fowler 1988; Williams & Andersen 1998) and these relationships could be incorporated into a model that could cover the entire spectrum of the injury process, from biopsychosocial factors related to the likelihood of experiencing an injury to the implications of these factors for the outcome of sports injury rehabilitation.

Conclusions

The proliferation of research investigating the psychological factors related to sports injury rehabilitation has demonstrated the importance of these factors to rehabilitation outcome. Descriptive, correlational, experimental, quantitative and qualitative research has shown that a variety of behavioural, cognitive, emotional and situational factors are involved in the sports injury rehabilitation process. Only through an adequate understanding of these factors and their interrelationships can well-designed interventions be proposed to enhance the rehabilitation process.
and, potentially, speed up or improve the outcome of sports injury rehabilitation.

Author note

Preparation of this chapter was supported in part by grant number R29 AR44484 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases. Its contents are solely the responsibility of the authors and do not represent the official views of the National Institute of Arthritis and Musculoskeletal and Skin Diseases.

References


PART 4

CLINICAL REHABILITATION INTERVENTIONS
Chapter 9

Pharmacological Agents and Acupuncture in Rehabilitation

JULIE K. SILVER AND JOSEPH AUDETTE

Introduction

Medications are a mainstay of treatment in the injured athlete—both for their pain relief and healing properties. Yet, despite the widespread use and acceptance of many medications, more often than not there is little research to endorse their use. In this chapter, we will focus on the primary medication classes that are used in sports medicine, especially during the first stage of rehabilitation. However, the reader is cautioned to recognize that the evidence that we rely so heavily on is sparse in this area. Therefore, it is recommended that all medications in athletes be used judiciously with a distinct regard for the risks and side effects as well as the potential benefits, which include pain relief and early return to play.

Non-steroidal anti-inflammatory drugs

A brief history

Inflammation has been around longer than doctors have known how to treat it. But since the very first physicians began to record what makes up the basis of our current medical knowledge, they have described inflammation. In the oldest known medical textbook, the Edwin Smith papyrus that was written by Egyptian healers 17 centuries before the birth of Christ, the body’s reaction to injury was called ‘shememet’. The term shememet is always followed by a hieroglyph for ‘fire’ that symbolizes the hallmarks of acute inflammation (Vertosick 2000). One thousand years later, the Greek physician Hippocrates coined the term ‘phlegmone’ which means ‘the burning thing’. Then, in the first century after the death of Christ, the Roman author Cornelius Celsus provided what is the basis for our description of inflammation today: rubor et tumour cum calore et dolore (redness and swelling with heat and pain) (Vertosick 2000).

It is not surprising that since our ancestral healers were so cognizant of inflammation that they would also have found ways to treat it. Indeed, ancient Egyptians and Assyrians used willow extract to reduce the redness and pain of inflamed joints; however, the first description of salicylate therapy did not come until several thousand years later. One historian notes that the ‘story begins in modern times with a letter from the Reverend Edward Stone of Chipping Norton to the Royal Society in 1763. He wrote: “There is a bark of an English tree which I have found by experience to be a powerful astringent and very efficacious in curing anguish and intermitting disorders”’ (Wall 2000).

Although there were various historical attempts to use willow bark and its derivatives, it was not until Hermann Kolbe, a professor of chemistry at Marburg University, succeeded in synthesizing salicylic acid (the precursor to modern aspirin) that it became widely used. In what is described as ‘a strange twist of history’ (Vertosick 2000), one patient who was disgusted with Kolbe’s concoction of salicylic acid that caused severe
gastrointestinal upset, approached his son, a chemist for the German manufacturer of chemical dyes, Fredrich Bayer & Company, and together they synthesized acetylsalicylic acid. They called this new drug Aspirin. Today there are more than 30 000 tons of aspirin sold worldwide, and aspirin and its cohorts in the non-steroidal anti-inflammatory drug (NSAID) class are universally used to control both pain and inflammation.

**Chemical properties and mechanism of action**

Today there are dozens of NSAIDs available commercially, which all have a common mechanism of action (Huff & Prentice 1999). Some of the more common NSAIDs are listed in Table 9.1. NSAIDs are widely used as first-line agents in sports injuries, particularly in soft tissue injuries that involve muscles, tendons and ligaments. These medications are used primarily for two reasons: (i) to decrease inflammation, and (ii) to reduce pain. NSAIDs also have antipyretic effects, but this is not usually a factor in their use for sports injuries. Because some of the NSAIDs are easily available without a prescription, they are often used haphazardly and improperly. Moreover, the literature has not clearly defined when NSAIDs should be used, what dose is most appropriate and for what period of time they should be used after an injury. Thus, even when they are prescribed by skilled clinicians, NSAIDs may cause more harm than good.

Inflammation is the normal response in tissue to any trauma—whether it is acute (macrotrauma) or chronic (microtrauma, generally due to repetitive motion). The purpose of the inflammatory response (see Chapter 2 for a more detailed discussion) is to contain the injury, remove the irreparably injured (necrotic) tissue and restore function by re-establishing structural integrity. The acute inflammatory response occurs most significantly in the first 48 h after injury. During this period of time, the inflammatory response is mediated by local vasoactive products such as histamine, bradykinin and serotonin, which are

### Table 9.1 Dosage of currently available NSAIDs.

<table>
<thead>
<tr>
<th>Generic name</th>
<th>Common dose (mg)</th>
<th>Usual dosing frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>325</td>
<td>Q 2–4 h</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>100</td>
<td>BID</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>75</td>
<td>BID</td>
</tr>
<tr>
<td>Diflunisal</td>
<td>500</td>
<td>BID</td>
</tr>
<tr>
<td>Etodolac</td>
<td>400</td>
<td>BID</td>
</tr>
<tr>
<td>Fenoprofen</td>
<td>600</td>
<td>QID</td>
</tr>
<tr>
<td>Flurbiprofen</td>
<td>100</td>
<td>TID</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>800</td>
<td>QID</td>
</tr>
<tr>
<td>Indomethacin</td>
<td>25–50</td>
<td>TID</td>
</tr>
<tr>
<td>Ketoprofen</td>
<td>75</td>
<td>TID</td>
</tr>
<tr>
<td>Ketorolac</td>
<td>10</td>
<td>QID</td>
</tr>
<tr>
<td>Meclomenamate</td>
<td>100</td>
<td>TID</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>500</td>
<td>2 QD</td>
</tr>
<tr>
<td>Naproxen</td>
<td>500</td>
<td>BID</td>
</tr>
<tr>
<td>Oxaprozin</td>
<td>600</td>
<td>2 QD</td>
</tr>
<tr>
<td>Piroxicam</td>
<td>20</td>
<td>QD</td>
</tr>
<tr>
<td>Rofecoxib</td>
<td>25</td>
<td>QD</td>
</tr>
<tr>
<td>Salicylsalicylic acid</td>
<td>750</td>
<td>QID</td>
</tr>
<tr>
<td>Sodium salicylate</td>
<td>650</td>
<td>Q 4 h</td>
</tr>
<tr>
<td>Sulindac</td>
<td>200</td>
<td>BID</td>
</tr>
<tr>
<td>Tolmetin</td>
<td>400</td>
<td>TID</td>
</tr>
</tbody>
</table>

BID, twice a day; Q 2–4 h, every 2–4 h; QD, each day; 2 QD, two once a day; Q 4 h, every 4 h; QID, four times a day; TID, three times a day.
released from mast cells in order to increase blood flow and permeability. Their effects predominate in the first hour after injury.

The acute inflammatory response is then partially maintained by a class of mediators called the eicosanoids, which leads to the formation of prostaglandins and leukotrienes. The primary focus of medications to control inflammation (in particular NSAIDs) has been to control prostaglandin synthesis, which is in large part responsible for the increased vascular permeability and vasodilatation at the site of injury. What happens in chronic injuries or those without a significant initial inflammatory phase is not well elucidated. For example, lateral epicondylitis or ‘tennis elbow’ was long thought to be a ‘tendinitis’ that was caused by an inflammatory condition. Not surprisingly, NSAIDs have been a first-line treatment for this condition. However, more recent studies reveal that microscopically the process is more consistent with a tendinosis—demonstrating angi-o-fibroblastic hyperplasia, hyaline degeneration, fibrinoid necrosis and granulation tissue with very few inflammatory cells (Nirschl 1992).

After the initial 24–48 h, the cellular inflammatory response begins the ‘clean-up’ process. This marks the beginning of the reparatory phase. The cells most involved in this process are macrophages and neutrophils that are responsible for digesting and clearing away unusable products in order to promote healthy tissue healing. It is in this phase that much of the controversy over NSAIDs occurs. The question that remains unanswered to date is whether NSAIDs retard this important phase of healing and whether this in fact might delay or impair the healing process (Leadbetter 1990).

Since the inflammatory response consists of vasodilatation with extravasation of blood that carries blood cells and other products into the area of injury, and this initial response is followed by the recruitment of inflammatory cells such as leucocytes and macrophages, then part of this process involves clearing the injured area of debris such as necrotic muscle tissue and disrupted connective tissues (Almekinders 1993). If the clean-up process is halted or impaired, it is not clear whether the healing process will suffer.

There are studies that have tried to address this, although the answer is not yet clear. For example, in one study 24 white rabbits sustained injury to the medial collateral ligament of one hindleg (Moorman et al. 1999). The rabbits were treated orally, twice daily, with a 2-week course of either ibuprofen or placebo. The ligaments were tested and there was no statistically significant difference in the values of the mechanical properties of the ligaments in the rabbits treated with ibuprofen versus placebo. The authors concluded that, under the conditions of the study, there was no deleterious effect of a short course of ibuprofen on the mechanical properties of medial collateral ligaments. However, in contradistinction to this study, another study was done on male rats where they underwent surgical transection of the medial collateral ligament (Elder et al. 2001). Postoperatively, half the rats were treated with celecoxib and the others were not. This study found that ligaments in the celecoxib-treated rats had a 32% lower load to failure than the untreated ligaments. Needless to say a reduced ligamentous strength is an unacceptable outcome in athletes eager to return to training and competition as early as possible.

Additionally, NSAIDs are known to affect the clotting process, and it is not clear in acute injury what impact this has on the tissues and the healing process. Unfortunately, studies have not been done to clearly delineate: (i) whether NSAIDs truly help to heal sports injuries or merely act as suppressers of the initial immune response as well as pain modulators; (ii) if they do help initially, then at what point do they become ineffective or even detrimental; and (iii) if they help for all injuries or just injuries in which one would anticipate a lot of inflammation such as an acute contusion or sprain/strain injury. There is a paucity of literature that illustrates when and how to use NSAIDs for sports-related injuries.

Regardless of the injury, NSAIDs have two main functions in the treatment of sports injuries: (i) to modify the inflammatory process; and (ii) to provide analgesia (Huff & Prentice 1999). The primary mechanism of action of NSAIDs in the inflammatory process is to inhibit prostaglandin
production. Prostaglandins are a group of fatty acids that help to mediate the inflammatory response. A decrease in prostaglandin synthesis through the inhibition of cyclooxygenase (COX) is believed to significantly inhibit the inflammatory response (Vane 1971). There are at least two forms of COX inhibitors present in humans (Huskisson et al. 1973; Polisson 1996). These enzymes act differently in the body and their actions in large part determine the side effects of NSAIDs. There has been increasing interest in NSAIDs that selectively promote COX-1 while inhibiting COX-2 (DeWitt et al. 1993). Examples of NSAIDs that selectively inhibit COX-2 are celecoxib and rofecoxib (see Table 9.1 for dosing).

This fairly straightforward explanation of how NSAIDs work is probably more complex in reality and has been challenged by a number of studies. Some investigators have suggested that NSAIDs may act directly on the inflammatory cells (Wahl et al. 1977; Ceuppens et al. 1986) and that perhaps some NSAIDs are fairly weak inhibitors of prostaglandin synthesis but appear to have a more pronounced analgesic effect (McCormack & Brune 1991). It is interesting to note that aspirin is the only NSAID that irreversibly inhibits COX; the other effects of other NSAIDs are reversible.

The mechanism by which NSAIDs produce analgesia appears to be multifactorial. Aspirin can interfere with the transmission of painful impulses in the thalamus (Moncada & Vane 1979). Blocking the inflammatory response is also probably a factor in decreasing pain. Both effects on pain and inflammation are thought to be due to the blocking of proinflammatory prostaglandins in the soft tissues.

Despite numerous studies elaborating on the mechanism of action of NSAIDs, it is not entirely clear whether they influence the outcome after sports injuries. Clearly, they attenuate the classic inflammatory response that consists of pain (dolor), heat (calor), redness (rubor), swelling (tumour) and loss of function (functio laesa) (Leadbetter 1990). But the degree to which they affect healing is not clear. In fact, they may actually slow healing in some instances. Moreover, in some injuries in which there is a marked inflammatory response, such as an acute contusion or sprain/strain injury, the effect of NSAIDs (either good or bad) may be more significant than in injuries where there is less of an inflammatory response (e.g. chronic repetitive strain injuries or delayed-onset muscle soreness, DOMS). Yet, despite the fact that compelling studies, which would support the use of NSAIDs in athletes, are lacking, there are studies that suggest a definite role for their use. For example, in one study where 364 Australian Army recruits with ankle sprains sustained during training were treated with placebo or piroxicam, the latter group had less pain and were able to resume training more rapidly than the placebo group (Slatyer et al. 1997). Two other studies revealed that NSAIDs in young adult males and in healthy older individuals attenuated the exercise-induced inflammation, strength loss and soreness associated with eccentric exercise (DOMS) (Dudley et al. 1997; Baldwin et al. 2001).

Dosing

Table 9.1 lists the common NSAIDs and their doses. When taken orally, NSAIDs typically reach steady-state concentrations in the serum after three to five half-lives (Stankus 1999). The clearance of these agents (usually via the liver) is variable, which accounts for the wide spectrum of elimination times. Thus, typically NSAIDs are prescribed for 2–3 weeks and monitored for their effectiveness. The use of multiple NSAIDs has not been shown to be more effective than monotherapy and increases the risk of adverse side effects (Stankus 1999). However, NSAIDs can be used safely with other analgesics if pain control is an issue (e.g. acetaminophen or opioids).

Topical preparations of NSAIDs have been shown to have a more reduced blood concentration than after oral or intramuscular administration (Doogan 1989; Heynemann 1995; Dominkus et al. 1996). However, they do appear to reach their target tissues and have been found in muscle and subcutaneous tissue after topical use (Dominkus
et al. 1996). Topical preparations of NSAIDs are gaining favour in sports medicine due to a number of promising studies that suggest objective and subjective improvement of symptoms (Akermark & Forsskahl 1990; Thorling et al. 1990; Russell 1991; Airaksinen et al. 1993).

**Side effects**

The side effects of NSAIDs result primarily from the inhibition of prostaglandin synthesis and occur in the following systems (in order of decreasing relative frequency): (i) gastrointestinal (e.g. gastritis or ulceration); (ii) renal (interstitial nephritis); (iii) dermatological (rash); and (iv) central nervous system (CNS) changes (Mortensen 1989).

By far the most notable side effect of NSAIDs is the toxicity to the gastrointestinal tract. Typical symptoms of gastric toxicity include heartburn or dyspepsia, gastritis and, potentially, ulceration. One-third to one-half of all patients who die of ulcer-related complications have recently been on NSAID therapy (Hollander 1994). Studies have demonstrated a markedly increased (4–30-fold) risk of developing ulcer disease associated with NSAID use (Soll et al. 1991; Griffith et al. 1991). Although the risk of gastrointestinal side effects is significant, it is important to note that most of the literature regarding this topic does not reflect its use in young, healthy athletes. Both topical NSAIDs and the oral selective COX-2 inhibitors may be safer alternatives regarding gastrointestinal toxicity than traditional oral NSAIDs (Dominkus et al. 1996; DeWitt et al. 1993). Additionally, medications used concurrently with NSAIDs, such as omeprazole, may provide some protection (Goodman & Simon 1994).

Renal toxicity is also an important consideration when using NSAIDs. Although serious renal complications are uncommon, there are examples of severe renal toxicity that have resulted from NSAID use. In one case report a 20-year-old female athlete developed end-stage renal failure that was attributed to ‘regular use of anti-inflammatory analgesic medication for minor sports injuries’ (Griffiths 1992). This occurred when she was in a stressful situation in a different country and was playing tennis for 4 h a day in a hot climate. Thus, it is likely that dehydration also contributed to her condition. There is also some evidence to suggest that in runners NSAIDs may be harmful due to a reduction in renal blood flow (Walker et al. 1994). It is important to remember that renal blood flow is reduced during exercise in healthy athletes and that proteinuria and haematuria have been reported after long-distance running and cycling (Eichner 1990; Mittleman & Zambraski 1992). Future studies are needed to investigate the possible interactions of exercise and NSAIDs.

Rashes have been noted with both oral and topical NSAIDs. Typically these are benign urticarial rashes that resolve when the medication is discontinued. But more severe dermatological rashes have been described, e.g. Stevens–Johnson syndrome and erythema multiforme (Stankus 1999). It is important to note that the triad of nasal polyposis, asthma and rhinitis may be an indication of increased risk of NSAID-induced hypersensitivity reaction (Stankus 1999).

Central nervous system side effects are uncommon but may include headache, tinnitus, depression, aseptic meningitis, mental status changes and coma (Stankus 1999). NSAIDs may alter blood pressure control or decrease the effectiveness of antihypertensive medications. NSAIDs also impair the normal function of platelets and may increase the bleeding time. Of note is that NSAIDs are not recommended in the pregnant athlete.

**Steroidal medications**

**A brief history and forms of administration**

Both corticosteroids and anabolic steroids have been used in the treatment of sports injuries. However, since the International Olympic Committee (IOC) (see doping issues, p. 200) and most other agencies governing the use of drugs in athletes ban anabolic steroids, this section will be primarily devoted to corticosteroid use—both
oral and injectable forms. Local anaesthetic medications that are often used in conjunction with corticosteroid injections (but may be used alone) will also be covered. The relevant restrictions on the use of local anaesthetics and glucocorticoids are discussed in the section on doping (p. 200).

The use of corticosteroids has a much more recent history than that of NSAIDs. It was not until the 1920s that Dr Philip Hench noted that patients with hypoadrenalism had many of the same symptoms as people with rheumatoid arthritis. Dr Hench concluded that rheumatoid arthritis must have a component of adrenal hormone insufficiency and could be cured with hormone replacement. Hench’s theory could not be tested until the 1930s when pure preparations of adrenal hormones became available. However, the hormone product that came from human adrenal glands (called compound E) was still not readily available. Years later, in 1948, Hench conducted the first clinical trial and the results were just as he suspected—patients with rheumatoid arthritis reacted miraculously to cortisol (compound E). Hench and his co-worker won the Nobel prize for medicine in 1950 and his Nobel address was titled, ‘The reversibility of certain rheumatic and non-rheumatic conditions by the use of cortisone’ (Vertosick 2000).

Although this finding was an important one, it was not what many people had hoped—a cure for rheumatoid arthritis and other medical conditions that involve inflammation. In fact, further studies revealed that cortisol had very serious side effects and could not be taken in large quantities for long periods of time. It is now common knowledge in the medical community and even in the public domain that corticosteroids produce significant deleterious side effects when taken orally.

In the injectable forms, although there are fewer systemic side effects, there are some potential hazards that mandate that these medications be prescribed with caution. It is critical that steroids used in injectable forms are properly delivered to the target site. Imaging techniques such as fluoroscopy are sometimes used to help guide the practitioner (Micheli & Solomon 1997). Injections also have the advantage of being both diagnostic and therapeutic. Favourable results are common with injectable steroids (Kapetanos 1982; Wiggins et al. 1994; Holt et al. 1995). Iontophoresis (using electrical stimulation) and phonophoresis (using ultrasound) are other methods of delivering corticosteroids locally and may prove to have some benefit (Franklin et al. 1995; Breen 1996).

Chemical properties and mechanism of action

The exact mechanism by which glucocorticoids work to mitigate inflammation is not entirely understood. Glucocorticoids are adrenocortical steroids that occur naturally and can be manufactured synthetically (Physicians’ Desk Reference 2002). These drugs are readily absorbed from the gastrointestinal tract. Prednisone is the most commonly prescribed synthetic oral glucocorticoid and will be the model for this discussion. In athletes, prednisone and the other glucocorticoids are prescribed primarily for their powerful anti-inflammatory properties.

Anabolic steroids are synthetic derivatives of the hormone testosterone. The actions of anabolic steroids are therefore similar to that of testosterone. In spite of the widespread use of adrenergic steroids by bodybuilders and other athletes, the effects of these drugs are poorly understood. Testosterone is known to increase muscle mass, but it is not known whether testosterone truly improves physical function and health-related quality of life (Bhasin et al. 2001). Anabolic steroids are banned in Olympic competitions. However, in clinical medicine there are some legitimate uses for these drugs including the promotion of weight gain after weight loss following extensive surgery, burns or severe trauma, to offset the side effects associated with prolonged corticosteroid use and for the relief of bone pain accompanying osteoporosis (Physicians’ Desk Reference 2002). More recently, there has been increased interest in the effect of male hormones and the synthetic derivatives on the healing of skeletal muscle (Beiner et al. 1999; Bhasin et al. 2001).
Dosing

Systemic corticosteroids are dosed in a variety of ways and the literature does not support a ‘best way’ approach. Often in an acute injury, corticosteroids are prescribed in a weaning fashion with the highest dose taken on the first few days and then tapering off the medication altogether within 1–2 weeks. There are commercially available ‘dose packs’ that are used for just this purpose. Longer term steroid use typically involves a more constant dose (e.g. 5 or 10 mg of prednisone daily), but is really practitioner-dependent. This approach is rarely used, if ever, in sports injuries and the systemic use of glucocorticoids is banned by the IOC (see doping issues, p. 200).

Injections can be done with corticosteroid alone or mixed with a local anaesthetic, at a number of sites (Table 9.2). In joints or bursae, aspiration of fluid may be done prior to injecting the medication. The dose and amount of injected material

<table>
<thead>
<tr>
<th>Injection sites</th>
<th>Needle</th>
<th>Volume (cm³)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joints</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulnocarpal</td>
<td>No. 25, 1.5 in</td>
<td>1–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Radiocarpal</td>
<td>No. 25, 1.5 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 2</td>
</tr>
<tr>
<td>Carpometacarpal</td>
<td>No. 25, 1.5 in</td>
<td>7–10</td>
<td>Steroid volume to anaesthetic ratio 1 : 6–9</td>
</tr>
<tr>
<td>Elbow</td>
<td>No. 25, 1.5 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 2</td>
</tr>
<tr>
<td>Shoulder</td>
<td>No. 25, 1.5 in</td>
<td>7–10</td>
<td>Steroid volume to anaesthetic ratio 1 : 6–9</td>
</tr>
<tr>
<td>Knee</td>
<td>No. 25, 1.5 in</td>
<td>3–5</td>
<td>Steroid volume to anaesthetic ratio 1 : 2–4</td>
</tr>
<tr>
<td>Sacroiliac joint</td>
<td>No. 27, 1.25 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Bone–tendon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med/lat epicondyle</td>
<td>No. 27, 1.25 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Plantar fascia</td>
<td>No. 27, 1.25 in</td>
<td>1–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Patellar tendon</td>
<td>No. 27, 1.25 in</td>
<td>3–5</td>
<td>Steroid volume to anaesthetic ratio 1 : 2–4</td>
</tr>
<tr>
<td>Achilles tendon</td>
<td>No. 27, 1.25 in</td>
<td>3–5</td>
<td>Steroid volume to anaesthetic ratio 1 : 2–4</td>
</tr>
<tr>
<td>Pubic symphysis</td>
<td>No. 27, 1.25 in</td>
<td>1–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Hamstrings</td>
<td>No. 27, 1.25 in</td>
<td>7–10</td>
<td>Steroid volume to anaesthetic ratio 1 : 6–9</td>
</tr>
<tr>
<td>Adductors</td>
<td>No. 27, 1.25 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 2</td>
</tr>
<tr>
<td>Tendon sheaths</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thumb extensor</td>
<td>No. 27, 1.25 in</td>
<td>1–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Finger flexors</td>
<td>No. 27, 1.25 in</td>
<td>3–5</td>
<td>Steroid volume to anaesthetic ratio 1 : 2–4</td>
</tr>
<tr>
<td>Posterior tibial</td>
<td>No. 27, 1.25 in</td>
<td>1–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Biceps (long head)</td>
<td>No. 27, 1.25 in</td>
<td>7–10</td>
<td>Steroid volume to anaesthetic ratio 1 : 6–9</td>
</tr>
<tr>
<td>Bursae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepatellar</td>
<td>No. 25, 1.5 in</td>
<td>2–3</td>
<td>If aspirating, will need No. 18–20 needle first</td>
</tr>
<tr>
<td>Pes anserine</td>
<td>No. 25, 1.5 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 1 : 1</td>
</tr>
<tr>
<td>Olecranon</td>
<td>No. 25, 1.5 in</td>
<td>7–10</td>
<td>Steroid volume to anaesthetic ratio 1 : 6–9</td>
</tr>
<tr>
<td>Subacromial</td>
<td>No. 25, 1.5 in</td>
<td>2–3</td>
<td>Steroid volume to anaesthetic ratio 2 : 1</td>
</tr>
<tr>
<td>Greater trochanter</td>
<td>No. 27, 1.25 in</td>
<td>2–5</td>
<td>Steroid volume to anaesthetic ratio 2 : 1</td>
</tr>
<tr>
<td>Perineural</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carpal tunnel</td>
<td>No. 27, 1.25 in</td>
<td>2–5</td>
<td></td>
</tr>
<tr>
<td>Tarsal tunnel</td>
<td>No. 27, 1.25 in</td>
<td>2–5</td>
<td></td>
</tr>
<tr>
<td>Suprascapular notch</td>
<td>No. 27, 1.25 in</td>
<td>2–5</td>
<td></td>
</tr>
<tr>
<td>Cubital tunnel</td>
<td>No. 27, 1.25 in</td>
<td>2–5</td>
<td></td>
</tr>
</tbody>
</table>
depends on the medication used and the structure that is being injected. There are no hard and fast rules on dosing, but some general guidelines advocated in the literature include: using 2.5–10 mg of prednisolone tebutate suspension in small joints (e.g. hand and foot), 10–25 mg in medium-sized joints (e.g. wrist and elbow) and 20–40 mg in large joints (e.g. knee and shoulder) (Swain & Kaplan 1995).

Side effects
The side effects associated with systemic corticosteroid use are many and often are associated with long-term use. Some of the more serious side effects include fluid and electrolyte disturbances, steroid myopathy, osteoporosis, aseptic necrosis of the femoral and humeral heads, peptic ulcer, pancreatitis, impaired wound healing, headaches, convulsions, suppression of growth in children, Cushingoid state, cataracts and glaucoma (Physicians’ Desk Reference 2002).

Injections carry the risk of injury to a blood vessel, nerve or other important structure. The deleterious effects on articular cartilage, tendons and the plantar fascia are somewhat disputed (Read & Motto 1992; Shrier et al. 1996), but there is the possibility of tendon or plantar fascial rupture following corticosteroid injection (Acevedo & Beskin 1998; Smith et al. 1999). An allergic reaction to the medication used, vasovagal syncope and infection are also important considerations (Table 9.3). Injectable steroids may also cause subcutaneous tissue atrophy and skin depigmentation, which are both usually just cosmetic concerns for the patient. When used by a skilled clinician, side effects are rare. Postinjection pain is rather common, however, and patients should be cautioned that they may experience a slight increase in pain for 1–2 days following the injection.

Muscle relaxants
The literature on the use of muscle relaxants in athletes is practically non-existent. It is mentioned here only because, anecdotally, physicians often do prescribe muscle relaxants for sprain/strain-type injuries. Whether muscle relaxants help or hinder healing is unclear. Equally unknown is whether muscle relaxants provide any pain relief effect or whether they allow athletes to return to their sport at an earlier date. What is known is that there are many different types of muscle relaxants that work by different pathways. Many of these have CNS side effects and at best cause fatigue or dry mouth and at worst can cause serious illness or even death (Linden et al. 1983; Kovac 1999; Olcina & Simonart 1999).

Opioids and inflammation
The use of opioids to treat athletic injuries is a controversial area. The reluctance of physicians to use these agents in the arena of sports injuries is due to a number of concerns that are both clinically and behaviourally motivated (Stanley & Weaver 1998). From a clinical point of view, opioids, as potent analgesics, may prevent the normal interpretation of pain signals by athletes and thereby allow too quick a resumption of aggressive training following injury. This could

<table>
<thead>
<tr>
<th>Table 9.3</th>
<th>Local and systemic complications with local injections of corticosteroids.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local</strong></td>
<td><strong>Systemic</strong></td>
</tr>
<tr>
<td>Subcutaneous atrophy</td>
<td>Transient hyperglycaemia in diabetics</td>
</tr>
<tr>
<td>Pigmentation abnormalities</td>
<td>Vasovagal symptoms with syncope</td>
</tr>
<tr>
<td>Tendon/ligament rupture</td>
<td>Cognitive effects, ‘steroid psychosis’</td>
</tr>
<tr>
<td>Accelerated joint destruction</td>
<td>Allergic reactions</td>
</tr>
<tr>
<td>Local sterile abscess</td>
<td>Systemic infection</td>
</tr>
<tr>
<td>Peripheral nerve injury</td>
<td>Suppression of pituitary–adrenal axis</td>
</tr>
<tr>
<td>Muscle necrosis/vascular injury</td>
<td>Avascular necrosis of hip</td>
</tr>
</tbody>
</table>
potentially lead to reinjury and further impairment. From a biobehavioural point of view, treating physicians must also be concerned about the inappropriate use of opioids to enhance performance. In addition, prolonged use of opioids to treat a more persistent injury can cause physical dependence and in rare cases psychological dependence, making cessation of the medication problematic.

Opioids or ‘narcotics’ are on the IOC list of prohibited substances (see p. 200) (Catlin & Murray 1996). However, given the need for more potent analgesics in treating acute injuries in sports, physicians lobbied the IOC to change its policy regarding opioids. As a result, in 1992, the IOC removed both codeine and dihydrocodeine bitartrate from the list of banned substances. It is interesting to note that codeine is a prodrug and is converted by the liver to morphine, which is the active form of the drug. However, approximately 8–10% of the population are genetically deficient in the hepatic enzyme needed to make this conversion and thus receive no analgesic benefit from codeine (Gourlay 1999). As we learn more about opioid physiology, especially with regard to pain and inflammation, it is likely that in the future selective use of opioids will be approved to treat athletic injuries.

There is increased appreciation of the role that both endogenous and exogenous opioids play in controlling not just pain, but inflammation following injury. The pharmacological actions of opioids have long been thought to reside exclusively within the CNS. Over the last 10 years, a growing understanding of the peripheral effect of opioid medications has developed. All three opioid receptor subtypes (mu, delta and kappa) have been isolated on peripheral axon terminals of small-diameter, unmyelinated sensory neurones (nociceptors or C-fibres). The opiate receptors are manufactured in the dorsal root ganglion and then transported both centrally into the dorsal horn and peripherally to the axon terminal. With inflammation, increased axonal transport of receptor proteins occurs with up-regulation of the expression of the receptors on the C-fibre nerve terminals (Stein et al. 2001).

In addition, inflammation causes the perineural barrier to become more permeable, presumably exposing more opioid receptors on the sensory axon to both exogenous and endogenous opioids (Coggeshall et al. 1997). This increased receptor expression is stimulated by the release of endogenous opioid peptides such as β-endorphin from CD4+ lymphocytes within inflamed tissue (Mousa et al. 2001). The source of the endogenous opioid seems to not be from pituitary secretion, but rather from circulating immune cells that accumulate at the site of acute inflammation. In particular, β-endorphin and met-enkaphalin have been found by radioimmunoassay in immunocytes at the site of injury in a rat paw model of acute inflammation (Cabot et al. 2001).

In animal models, the evidence suggests that part of the anti-inflammatory action of opioids is to reduce the release of proinflammatory mediators by the axon terminals. Peripheral nociceptors are now understood to not only signal tissue injury centrally, but they also act to release peptides in the periphery, such as substance P and calcitonin gene-related peptide which promote inflammatory cell activity at the site of injury. This axonal reflex to injury is referred to as neurogenic inflammation and is also believed to be responsible for tissue swelling and pain outside the area of direct tissue injury. Excessive release of these proinflammatory substances can lead to peripheral sensitization of sensory afferents, which is in part due to an increase in the expression of sodium channels (Fig. 9.1) (Julius & Basbaum 2001). This in turn leads clinically to hyperalgesia or an exaggerated response to minimally painful stimulation (Woolf & Salter 2000). By diminishing the release of these proinflammatory peptides from the C-fibre terminals, opioids may be important in limiting the volume and extent of such sensitization. In addition, opioid receptors have been found on immune cells suggesting a mechanism for the effect of morphine on suppressing lymphocyte proliferation and diminishing the production of various cytokines including interleukin-1β, an important immune modulator (House et al. 2001).
Clinical applications

Use of intra-articular, preservative-free morphine sulphate has become more prevalent after joint operations such as knee meniscectomy. The goal is to reduce postoperative pain and inflammation and speed recovery. With meniscectomy, the usual recovery time course can last up to 10 weeks before resumption of athletic competition. In general, the more pain and swelling following surgery or injury, the longer the time to recovery. In a recent study that had as outcome measures (i) time to being pain free, (ii) time before cessation of crutches, and (iii) time before return to work, the use of intra-articular morphine with bupivacaine allowed a more than 50% reduction in average time necessary to reach these endpoints when compared to intra-articular saline injection (Rasmussen et al. 1998). The addition of methylprednisolone to bupivacaine and free morphine sulphate led to even greater shortening of the duration before full recovery. Similar findings have been reported following ankle surgery in athletes (Rasmussen & Kehlet 2000). In comparing intra-articular morphine (3 mg in 3 cm³ saline) with dexamethasone (4 mg in 3 cm³ saline) in chronic arthritis patients, both produced a significant reduction of pain when compared with saline, and the morphine group had significantly lower synovial leucocyte counts when compared with the dexamethasone group (Stein et al. 1999).

In a review by Kalso et al. (1997), 33 randomized, controlled trials studying 1500 patients were found comparing intra-articular morphine following knee surgery with placebo treatments. Dose ranges of the morphine ranged from 0.5 to 5 mg and no additional benefit was seen with doses above 1 mg per knee. Overall, the studies suggest that there is a prolonged analgesic effect and reduction in consumption of pain medications. However, various methodological problems with the study designs make it difficult to conclude that the use of intra-articular morphine has a significant, clinically relevant effect on postsurgical outcomes.

Side effects

Side effects associated with short-term opioid use mainly involve cognition and the gastrointestinal system. The side effects for all of the commonly

---

**Fig. 9.1** Schematic diagram of the response of nociceptors to tissue injury. ASIC3, proton receptor; ATP, adenosine triphosphate; BK₂, bradykinin receptor 2; CB1, cannabinoid receptor 1; CGRP, calcitonin-gene related peptide; DEG, pressure receptors; DRG, dorsal root ganglion; ENaC, putative pressure receptor; H, protons; 5-HT, serotonin (5-hydroxy-tryptophan; 5-HT₃, serotonin receptor 3; NGF, nerve growth factor; PGF₂α, prostaglandin F-2 receptor; P₂X₃, purinergic receptor; TrkA, tyrosine kinase receptor A; VR1, vanilloid receptor 1; VRL-1, vanilloid receptor like protein. (From Julius & Basbaum 2001, with permission.)
prescribed opioid medications (Table 9.4) are essentially similar, although there is a great degree of variation in the side effects caused by various opioid agents for a particular individual. So whereas codeine may cause severe nausea and cognitive disorientation for a particular person, hydrocodone may not.

Cognitively, opioid-naïve individuals may feel disinhibited, with loss of cognitive acuity and mild to moderate sedation when first taking a potent narcotic. Usually these cognitive side effects dissipate with continued use. Associated with the cognitive side of sedation is respiratory depression, especially in the postoperative setting. This is associated with a shift in the responsiveness of the respiratory drive to the carbon dioxide concentration in the lungs, an important mechanism of respiratory regulation during exercise. Although pain is a potent stimulus to overcome this opioid effect, use of other agents to relieve pain, including spinally administered anaesthetics, can at times lead to the sudden relief of pain and the rapid onset of respiratory depression if the dose of opioid medication is not reduced. Other medications can act synergistically with opioids to cause respiratory depression, especially agents in the benzodiazepine and barbiturate class. Carisoprodol is often used as a muscle relaxant for acute pain and can be a particularly bad combination with opioid medications, as it metabolizes to meprobamate, which is a barbiturate.

Some individuals experience a mild euphoria with the consumption of an opioid medication. This response to opioids may be predictive of an individual being prone to develop addictive behaviours around the use of opioid medications; however, there is no well designed study to make a definitive statement regarding this issue. Of note is the fact that a small percentage of individuals will experience extreme dysphoria, to the point of depression, with the consumption of opioids. Often this will be accompanied by some degree of agitation, both of which resolve with cessation of the agent. Finally, prolonged use of opioid medications can cause reduced libido, but this is not usually a major issue with short-term use (Evans 1999).

Gastrointestinal side-effects are extremely common. Mild to severe nausea can occur, but the most common problem that uniformly affects individuals taking opioids is constipation. Appropriate measures to increase water and fibre intake can usually overcome this problem, but bowel obstruction can occur if left unattended. Less common problems include excessive sweating, pruritis and urinary retention. Extremely high doses of any opioid, but particularly found with

<table>
<thead>
<tr>
<th>Generic name</th>
<th>Morphine equivalent dose, oral (mg)</th>
<th>Starting dose</th>
<th>Usual dosing frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenazocine</td>
<td>50–200</td>
<td>50 mg</td>
<td>QID</td>
</tr>
<tr>
<td>(agonist/antagonist)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propoxyphene</td>
<td>65</td>
<td>50–100 mg</td>
<td>QID</td>
</tr>
<tr>
<td>Codeine</td>
<td>30–200</td>
<td>30–60 mg</td>
<td>QID</td>
</tr>
<tr>
<td>Morphine</td>
<td>5</td>
<td>15 mg</td>
<td>Q 4 h</td>
</tr>
<tr>
<td>Hydrocodone</td>
<td>10</td>
<td>5–7.5 mg</td>
<td>Q 4 h</td>
</tr>
<tr>
<td>Meperidine</td>
<td>200</td>
<td>50 mg</td>
<td>Q 4 h</td>
</tr>
<tr>
<td>Hydromorphone</td>
<td>2–4</td>
<td>1–2 mg</td>
<td>Q 3–4 h</td>
</tr>
<tr>
<td>Fentanyl patch</td>
<td></td>
<td>25 μg/h</td>
<td>72 h</td>
</tr>
<tr>
<td>Oxycodone</td>
<td>30</td>
<td>5–10 mg</td>
<td>Q 4 h</td>
</tr>
<tr>
<td>Methadone</td>
<td>8–10</td>
<td>2.5–5 mg</td>
<td>Q 8–12 h</td>
</tr>
<tr>
<td>Levorphanol</td>
<td>4</td>
<td>2 mg</td>
<td>Q 6–8 h</td>
</tr>
</tbody>
</table>

Q 3–4 h, every 3–4 h; Q 4 h, every 4 h; Q 8–12 h, every 8–12 h; Q 6–8 h, every 6–8 h; QID, four times a day.
meperidine, can lower seizure threshold and can occasionally cause cardiac arrhythmias. However, this is usually not important with short-term use in the setting of injury (Bowdle 1998). Often, the most serious physiological side effect in the outpatient setting from the use of high-dose opioids is hepatitis due to excessive acetaminophen intake found in the combination of short-acting opioids such as oxycodone and hydrocodone.

Finally, the issue of addiction versus tolerance must be raised when discussing the use of opioids in pain. Tolerance occurs with prolonged use of opioids uniformly in all patients and is due to physiological changes in the CNS. Clinically, this is expressed as the need to increase the dose of opioid medication to achieve continued analgesia. Addiction is a biobehavioural phenomenon and is described as a maladaptive, self-destructive activity to continue to obtain and seek more opioid medications despite progressive deterioration in function and social status. Surveys in the acute pain population, such as those with postoperative pain, show that it is extremely rare to see the onset of addictive behaviours with the appropriate use of opioid medications. In the chronic pain population, estimates of the incidence varies from 3.2 to 18.9% (Fishbain et al. 1992).

In the future, as we learn more about the effect of opioids on peripheral pain and the inflammatory transduction system of tissue injury, novel methods of treating the pain and inflammation associated with athletic injuries will probably include agents that act on these opioid receptors. Ideally, agents that do not have a significant CNS effect could be used to modulate the immunoinflammatory response to injury.

**Acupuncture in acute pain and inflammation**

Discussion of the peripheral effects of endogenous and exogenous opioids would be incomplete without some mention of the effect of acupuncture on pain and tissue injury. The use of acupuncture by athletes to treat acute injuries is very common in Asia and Eastern Europe and is increasingly becoming an option in many Western nations. At the Winter Olympics in Japan in 1998, international exposure came when the acupuncturist in Nagano offered free treatments to Olympic athletes and officials, emphasizing that it is a drug-free way to treat injuries. Even more stunning was the near miraculous recovery in response to acupuncture by the Austrian, Hermann Maier. Maier won gold medals in the giant slalom and super G, 3 days following a dramatic fall and injury that occurred during the downhill competition. Maier mentioned to the press that the use of acupuncture to treat his shoulder and knee injuries following the fall helped him to recover so quickly.

Over the last 30 years, a great deal of scientific evidence has accumulated to substantiate that acupuncture stimulation (AP) and electroacupuncture stimulation (EA) have physiological effects that strongly influence the neurohumoral systems that modulate pain. The evidence for the release of endogenous opioids with AP and EA goes back to the seminal work of Pomeranz in animals and Mayer in humans in the early 1970s (Pomeranz & Chiu 1976; Mayer et al. 1977). We now know that acupuncture causes the release of endorphins and enkephalins in the CNS and that these neuropeptides play a significant role in its analgesic efficacy. There is growing evidence that the descending inhibitory control system also plays a role in acupuncture analgesia. This involves the activation of serotonergic neurones in the midbrain, that in turn act to inhibit the transmission of nociceptive information at the level of the dorsal horn (Fig. 9.2). Release of 5-hydroxytryptamine (5-HTP) in the raphe nucleus of the midbrain has been shown with both EA and AP, and acupuncture analgesia is attenuated with the injection of a serotonin antagonist such as parachlorophenylalanine (Debreceni 1993).

Acupuncture has been shown in animal models to strongly influence the pituitary–hypothalamic system as well. The arcuate nucleus of the ventromedial hypothalamus contains the β-endorphin-producing cells and lesions of this nucleus abolishes acupuncture analgesia in rats (Debreceni 1993). The hypothalamus secretes
β-endorphin into the blood, consistent with the elevated blood levels of endorphin found with both AP and EA. Concurrent with the release of β-endorphin, is the secretion of adrenocorticotropic hormone with acupuncture, which in turn stimulates adrenal secretion of cortisol causing a general anti-inflammatory effect (Pomeranz 1998). Given the previous discussion relative to the presence of peripheral opioid receptors on C-fibres in inflammatory conditions, the elevation of systemic β-endorphin with acupuncture may exert both an analgesic and an anti-inflammatory effect in sports injuries.

Recent studies indicate that acupuncture also influences the release of immune-modulating cytokines from the hypothalamus such as interleukin 1β and 6 during experimental models of fever (Son et al. 2002). The rise of blood β-endorphin with acupuncture also influences peripheral cytokine production in the spleen by causing increases in interferon-γ and reduction of natural killer cell activity (Yu et al. 1998). These early data suggest that acupuncture has physiological effects that go far beyond the release of endorphins in the CNS, and has important immune-modulating effects that may prove to be important in the recovery from injury.

Clinical applications

The use of acupuncture for acute injuries has not been studied extensively in athletes. However, a recent study of rotator cuff tendinitis in subjects with sports-related injuries found that true acupuncture needling was significantly superior when looking at both its analgesic effects as well as strength, range of motion and functional scores when compared to a placebo needling control group (Kleinhenz et al. 1999). Treatment consisted of using a combination of points in the feet and hands, based on Chinese acupuncture principles, together with meridian-based local points around the shoulder based on tenderness. Subjects received two 22 min treatments per week for 4 weeks and had follow-up assessments at the end of the treatment period and again at 4 months. Acupuncture has also shown efficacy in providing acute pain relief with a single 20 min treatment for chronic lateral epicondylitis (tennis elbow) when compared with placebo. A single treatment was given using only one point near the fibular head (gallbladder 34) on the leg ipsilateral to the elbow pain (Molsberger & Hille 1994). Both of these studies used a placebo needle device for the control group, where a blunt
needle presses on the point, but does not penetrate the skin.

In a randomized trial of acupuncture for osteoarthritis of the knee, significant improvement in scores of pain and function were noted when compared with standard treatment on oral NSAIDs (Berman et al. 1999). In this study, subjects received 20 min sessions of electroacupuncture at 2.4–4 Hz for 8 weeks with outcome assessments at 4, 8 and 12 weeks. Finally, in a meta-analysis of randomized controlled trials of acupuncture for low back pain, it was concluded that acupuncture is superior to the various control treatments, but insufficient evidence was available to state whether it was better than placebo needling methods (Ernst & White 1998).

In conclusion, acupuncture appears to be a safe and potentially effective method of treating the pain and inflammation of sports-related injuries. There is sufficient evidence to suggest a physiological mechanism of action that involves neurohumoral processes that have both a central and peripheral effect. Better designed and more specific studies are needed to assess the clinical efficacy of acupuncture in various sports injuries. It is also important to understand how acupuncture might affect human performance in athletic competitions.

### Doping issues

It is very important for every sports medicine practitioner to become familiar with the list of prohibited substances of the IOC (Table 9.5). Relevant to this chapter are sections IB, IC, IIIC and IIID.

Narcotics such as diamorphine (heroin), methadone, morphine and related substances are prohibited in Olympic sports competitions. On the other hand, codeine, dextromethorphan, dextropropoxyphene, dihydrocodeine, diphenoxylate, ethylmorphine, propoxyphene and tramadol are permitted. The sports physician must know the content of various medications to avoid using prohibited substances. Clearly, the doping rules do not apply in an emergency when the life of an athlete may be at risk.

<table>
<thead>
<tr>
<th>Table 9.5 List of prohibited substances and methods in sports according to the International Olympic Committee.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I Prohibited classes of substances</strong></td>
</tr>
<tr>
<td>A. Stimulants</td>
</tr>
<tr>
<td>B. Narcotics</td>
</tr>
<tr>
<td>C. Anabolic agents</td>
</tr>
<tr>
<td>1. Anabolic androgenic steroids</td>
</tr>
<tr>
<td>2. β2-agonists</td>
</tr>
<tr>
<td>D. Diuretics</td>
</tr>
<tr>
<td>E. Peptide hormones, mimetics and analogues</td>
</tr>
<tr>
<td><strong>II Prohibited methods</strong></td>
</tr>
<tr>
<td>A. Blood doping</td>
</tr>
<tr>
<td>B. Artificial oxygen carriers or plasma expanders</td>
</tr>
<tr>
<td>C. Pharmacological, chemical and physical manipulation</td>
</tr>
<tr>
<td><strong>III Classes of prohibited substances in certain circumstances</strong></td>
</tr>
<tr>
<td>A. Alcohol</td>
</tr>
<tr>
<td>B. Cannabinoids</td>
</tr>
<tr>
<td>C. Local anaesthetics</td>
</tr>
<tr>
<td>D. Glucocorticosteroids</td>
</tr>
<tr>
<td>E. β-blockers</td>
</tr>
</tbody>
</table>

Anabolic androgenic steroids are prohibited in Olympic sports and there is no indication or exemption to this rule. Since the β2-agonists have anabolic properties they have been included in the list of prohibited substances. Some β2-agonists like formoterol, salbutamol, salmeterol and terbutaline are permitted by inhaler only to prevent and/or treat asthma and exercise-induced asthma. Written notification by a respiratory or team physician that the athlete has asthma and/or exercise-induced asthma, is necessary to the relevant medical authority prior to competition. At the Olympic Games, athletes who request permission to inhale a permitted β-agonist will be assessed by an independent medical panel.

Local anaesthetics are permitted under certain conditions. Bupivacaine, lidocaine, mepivacaine, procaine and related substances can be used, but not cocaine. Vasoconstrictor agents may be used in conjunction with local anaesthetics but only local or intra-articular injections may be administered, and only when medically justified. Notification of administration may be necessary where the rules of a responsible authority so provide.
The systemic use of glucocorticoids is prohibited when administered orally, rectally or by intravenous or intramuscular injection. When medically necessary, local and intra-articular injections of glucocorticoids are permitted. Notification of administration may be necessary where the rules of a responsible authority so provide.

References


increased risk of peptic ulcer disease in elderly patients. *Annals of Internal Medicine* 114, 257–263.


Chapter 10

Physical Modalities and Pain Management

JOEL M. PRESS, CHRISTOPHER T. PLASTARAS AND STEVEN L. WIESNER

Introduction

This chapter provides the reader with the fundamentals of various modalities used in the treatment of sports-related injuries. Heat, cold and the use of electrical stimulation may be useful adjunctive components of the rehabilitation programme. By understanding the physiological basis of these modalities, a safe and appropriate treatment choice can be made. One must remember, however, that the most effective modality used will ultimately depend upon the patient’s individualized and subjective response to treatment. Lastly, and most importantly, the use of therapeutic modalities is only one component of a comprehensive rehabilitation programme.

Cryotherapy (cold therapy)

As with heat, in order to safely apply and more effectively utilize cryotherapy in the treatment of sports-related injuries, an understanding of the physiological effects is necessary. In comparison with heat, cold produces vasoconstriction with vasodilatation following reflexively; it causes decreased local metabolism; and it minimizes enzymatic activity and the subsequent demand for oxygen (Lehmann & de Lateur 1982a; Ork 1982). Cold diminishes muscle spindle activity as well as slowing nerve conduction velocity. As a result, cold is commonly used for pain control and decreasing muscle spasticity and muscle guarding (Chambers 1969; Mennell 1975; McMaster 1982). Cryotherapy, in comparison to therapeutic heat, has the opposite effect on collagen extensibility. Cold increases connective tissue stiffness and muscle viscosity, thereby diminishing flexibility.

Based on these various physiological effects, cryotherapy is most commonly used during the first 48 h of acute musculoskeletal injuries such as sprains, strains and contusions (Grana et al. 1986). Use beyond the acute phase is justified for continued pain control, muscle re-education and control of swelling when utilized with compression (Quillen & Rouillier 1982; Sloan et al. 1988). Contraindications for cryotherapy include ischaemia, cold intolerance, Raynaud’s phenomenon, cold allergy, inability to communicate and insensate skin (Basford 1998). Care must be taken when using cold therapy over nerves due to the potential development of neuropraxia (conduction block). Recommendations to minimize this complication include limiting ice application to less than 30 min and protecting any peripheral nerves in the region (Drez et al. 1981).

Techniques of application

Ice packs, iced compression wraps, slushes, ice massage, ice whirlpools and vapocoolant sprays are some methods of cold application (Basford 1998). Regardless of the method used, there is a rapid drop in skin temperature with a delayed effect on muscle. This depends on the amount of overlying subcutaneous tissue, with maximum cooling occurring to a muscle depth of 1–2 cm (Halvorson 1989). Ice continues to be the safest
and most effective method of application. Care must be taken when using chemical or gel packs due to the poor control of temperature and risk of skin irritation should the envelope break and the chemical come in contact with the individual’s skin (Grant 1964; McMaster et al. 1978; Lehmann & de Lateur 1982a). Cryostretch and cryokinetics refer to the use of cryotherapy to facilitate joint movement. Decreasing pain and muscle guarding may lead to improved flexibility and muscle function (Roy & Irvin 1983; Halvorson 1989). An additional method of cryotherapy involves the use of vapocoolant sprays (fluori-methane and ethyl chloride), which provide very effective cutaneous local anaesthesia and are commonly used to treat myofascial trigger points. The use of cryotherapy in this context promotes normal muscle resting length by a ‘spray and stretch’ technique rather than by cooling the muscle itself (Mennell 1975). Fluori-Methane is less explosive, less flammable and produces less cooling than ethyl chloride (Travell & Simons 1983). Concerns, however, have been raised regarding the destruction of the ozone layer by the use of vapocoolant sprays, some of which are considered to be chlorofluorocarbons (Vallentyne & Vallentyne 1988; Simons et al. 1990).

Schaubel (1946) showed that the use of ice in 345 patients decreased the need for analgesic medication following assorted orthopaedic operations, as compared with controls. In a study using both clinical and basic science models, Ohkoshi et al. (1999) showed that cryotherapy is helpful with pain scores and analgesic use in postoperative anterior cruciate ligament reconstruction patients. Twenty-one patients undergoing reconstructive anterior cruciate ligament reconstruction were randomized into three groups: cryotherapy at 5°C, cryotherapy at 10°C and no cryotherapy for the first 48 postoperative hours. A temperature probe was placed intraoperatively in the suprapatellar pouch and the intercondylar notch through arthroscopic portals. The cryotherapy groups reached 120° of knee flexion 4 days sooner than controls. The 10°C group had significantly lower pain scores and analgesic use as compared with the control group. Both the cryotherapy groups showed a triphasic temperature curve with a low temperature phase occurring immediately postoperatively and lasting about 2 h; followed by a temperature rising phase; then finally a thermostatic phase. The control group, however, went immediately to a thermostatic phase. During the low temperature phase the cryotherapy groups’ suprapatellar pouch temperatures were significantly lower than the intercondylar notch temperatures—both sites being significantly lower than body temperature. Only the suprapatellar pouch temperature remained significantly lower than body temperature during the thermostatic phase. Although the numbers were not large in this interesting study, it helps link the objective intra-articular temperature changes to clinical outcomes (Fig. 10.1). Martin et al. (2001) also found reduced intra-articular knee temperatures with ice application with compression.

In a randomized, controlled study, Konrath et al. (1996) did not find significant differences in medication usage or range of motion in 100 postoperative anterior cruciate ligament reconstruction patients treated with cryotherapy; unfortunately, pain scores were not recorded. Levy and Marmar (1993) reported less swelling, less pain and better range of motion with the use of cryotherapy in patients following total knee arthroplasty.
Levy et al. (1997) studied the effect of cryotherapy on temperatures in the glenohumeral joint and the subacromial space following shoulder arthroscopy in 15 patients. They found no significant differences in temperatures as compared with controls; however, clinical information on pain scores was not reported. Pain scores were reported in a randomized, controlled study of cryotherapy in 50 postoperative shoulder patients done by Speer et al. (1996). The cryotherapy group reported decreased pain frequency and intensity, less need to use medication, better sleep, less swelling and less pain with shoulder movement. Karlsson et al. (1994) found significantly improved pain scores with cryotherapy following arthroscopic anterior cruciate ligament reconstruction. Similarly, Lessard et al. (1997) also showed decreased pain scores and analgesic use following arthroscopic knee surgery in a randomized, blinded, controlled study of 45 patients.

Cryotherapy has been shown to help patients with sprained ankles return to activity an average of 5.1 days more quickly (Basur et al. 1976). Hocutt (1981) found that early cryotherapy returned patients with sprained ankles back to activity an average of 15 days earlier when compared with late cryotherapy or heat therapy.

Paddon-Jones and Quigley (1997) and Yackzan et al. (1984) did not find cryotherapy to be effective in treating delayed onset muscle soreness. In a study of rats, Fu et al. (1997) found that post-endurance training cryotherapy may actually be deleterious by causing histological myofibril damage.

Merrick et al. (1999) found that 5 h of continuous cryotherapy after a crush injury to the triceps surae of rats reduced biochemical secondary injury compared with no treatment. Ho et al. (1994) showed that 20 min of cryotherapy on the knees was enough to reduce arterial blood flow, soft tissue blood flow and bone uptake. MacAuley (2001) reviewed 45 textbooks and found considerable variation in the recommended duration and frequency of the use of ice. Although there is no definitive recommendation regarding the duration of treatment, cryotherapy is typically used for a period of 20–30 min at a time. There is more risk of frostbite if ice is used directly on the skin. In a review of cryotherapy in sports medicine, Swenson et al. (1996) deemed cryotherapy as ‘effective and harmless’, as few complications are reported.

### Contrast baths

Contrast baths, which alternate the use of heat and cold, have been described as a form of ‘vascular exercise’ due to alternating dilatation and constriction of blood vessels. By alternating cycles of heat and cold, a hyperaemic response may be created, thereby improving circulation and assisting in the healing response. Specific indications would include improving range of motion, controlling swelling and providing pain control. Contraindications include those already discussed for therapeutic heat and cold, particularly active bleeding and vascular insufficiency. A protocol commonly used is as follows (Lehmann & de Lateur 1982a; Halvorson 1989):

1. The affected region is submerged in a warm bath of 38–43°C for 10 min.
2. This is followed by a cold bath of half ice/half water at 13–18°C for 1 min.
3. A warm bath is then used for 4 min.
4. This is followed by a cold bath for 1 min.
5. Steps 3 and 4 are then alternated for a total treatment cycle of 20–30 min.
6. The sequence ends with a cold bath.

Exercising the area may occur during the heating phase with rest during the cooling phase. A somewhat different protocol is described by Roy and Irvin (1983) in which the treatment begins and ends with cold immersion. However, it is not clear that this technique works better than cold alone.

### Therapeutic heat modalities

In order to better understand the appropriate role for heat therapy, one needs to be aware of the various modes available. Heat can be transferred to tissue by three methods.

1. **Conduction**: Direct heat transfer from one surface to another due to direct contact. This
pain management 207

is a form of superficial heat. Examples would include hydrocollator packs and paraffin baths. The penetration depth depends on the thickness of the subcutaneous tissue.

2 *Convection*: Heat transfer due to the movement of air or water across a body surface. This, too, is a form of superficial heat. Examples would include hydrotherapy and fluidotherapy.

3 *Conversion*: Transfer of heat due to a change in the form of energy. Examples of superficial heat conversion would include radiant energy such as that produced by infrared lamps. Deep heating, also known as diathermy, is due to conversion through the use of short waves, electromagnetic microwaves and ultrasound.

To best utilize therapeutic heat modalities, an understanding of the physiological effects is necessary (Table 10.1). Heating can create changes which are both local and distant, with the far-field effects being less pronounced. A consensual response may also be seen whereby heating one part of the body creates an increase in blood flow in other regions. As heat is applied to a body surface, circulatory changes occur which include vasodilatation of the arterioles and capillaries. As a result of increased metabolic tissue demands, there is a subsequent increase in blood flow with the arrival of various leucocytes, improved delivery of oxygen, increased capillary permeability and hyperaemia. Due to the increased blood flow, diffusion across membranes occurs more effectively leading to oedema formation, especially during the acute process (Cox et al. 1986).

Additional physiological effects include pain control through vasodilatation, which, when applied during the later stages of healing, leads to improved blood flow through the removal of pain-causing inflammatory substances such as bradykinins, prostaglandin and histamine substrates. Heat also acts directly on free nerve endings and provides muscle relaxation by decreasing the muscle spindle’s sensitivity to stretch via the gamma system (Lehmann & de Lateur 1990). Inhibitory pathways can be activated by the use of heat modalities with subsequent muscle relaxation. Central processes may also provide for sedation and decreased pain awareness. Thus, therapeutic heat assists in altering the pain–muscle guarding (spasm) cycle (Fountain et al. 1960).

One of the most useful therapeutic effects of heat includes improved collagen flexibility, especially when accompanied by prolonged stretching, as well as a subjective decrease in joint stiffness. Lehmann et al. (1970) showed that lengthening of the tendon and decreasing tendon tension occurred most effectively when the tendon was loaded in an elevated temperature bath of 45°C as compared with one at 25°C. Furthermore, only when stretch was applied in conjunction with heat did lengthening occur (Warren et al. 1971, 1976b). There are some studies, however, that do not support this. In a small study of 24 subjects, Taylor et al. (1995) showed that heat or cold modality made no significant difference in hamstring length when used in conjunction with a sustained hamstring stretch.

Based on the physiological effects of therapeutic heat, indications for heat modalities with specific attention to the injured athlete include pain, muscle spasm, contracture, bursitis and tenosynovitis. Contraindications for heat include peripheral vascular disease, bleeding diathesis, malignancy, acute trauma, sensory deficits or in patients that are unable to communicate about their sensation of pain (Basford 1998). Therapeutic heat, like other modalities, can often provide short-term relief, but there is little evidence to

Table 10.1 Physiological effects of heat.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasodilatation of arterioles and capillaries</td>
<td>Improved oxygen delivery to tissues</td>
</tr>
<tr>
<td>Increased diffusion across membranes</td>
<td>Oedema formation</td>
</tr>
<tr>
<td>Hyperaemia</td>
<td>Removal of inflammatory substances</td>
</tr>
<tr>
<td>Increased metabolic tissue demands</td>
<td>Decreases muscle spindle sensitivity, thereby encouraging muscle relaxation</td>
</tr>
</tbody>
</table>

PAIN MANAGEMENT 207
support long-lasting effects. Timm (1994) studied 250 subjects who had persistent low back pain following an L5 laminectomy in a randomized, controlled trial. The subjects were randomized into five groups including: (i) control; (ii) manipulation; (iii) simple home exercise programme; (iv) supervised exercise programme; and (v) physical agents, including hot packs, ultrasound and transcutaneous electrical nerve stimulation (TENS). The physical agent group did no better than the control group on the functional Oswestry scale, but was the most costly of all groups (US$1842 per subject). The exercise groups had the most improvement in the Oswestry disability scores and had fewer recurrences of low back pain; the simple home exercise programme was also the most economical (US$1392 per subject).

**Superficial heat modalities**

The common denominator of superficial heat modalities is direct heat penetration. Penetration is greatest within 0.5 cm from the skin surface, depending on the amount of adipose tissue (Lehmann et al. 1966; Michlovitz 1986). The more commonly used modes for sports rehabilitation are hydrocollator packs, whirlpools and contrast baths (see above). Other forms of superficial heat include infrared lamps, paraffin baths, fluidotherapy and moist air.

**Hydrocollator packs**

Hydrocollator packs serve to transfer heat via conduction. These packs come in three standard sizes and are heated in stainless steel containers containing water at temperatures between 65 and 90°C (Griffin & Karselis 1982). After appropriate heating, toweling is applied to the packs in order to minimize burning of the skin and to maintain heat insulation. The highest temperatures produced by the hydrocollator packs are at the skin surface. The pack is able to maintain heat for approximately 30 min with treatment sessions lasting 20–30 min.

Other heating packs are also available and include hot water bottles, electric heating pads and chemical packs. The major disadvantage of using these devices is limited temperature control.

**Hydrotherapy**

**Exercise in water**

Hydrotherapy is a term that can describe two distinct entities: warm water immersion or exercise performed in the water. Warm water immersion is discussed below.

A patient exercising in the water can get the therapeutic benefit of exercise while using the buoyancy principles of water to decrease the biomechanical stresses on the musculoskeletal system. The temperature of the water can be modified to fulfil individual needs. Patients with acute injuries and pregnant women are typically treated in cooler pools of 28°C whereas subacute or chronic injuries are treated in warmer temperatures of 33–34.5°C (Konlian 1999). Water exercises can also be used to cross-train patients who require weightbearing restrictions, such as those with stress fractures. Buoyancy-assisting devices can be used to allow patients to run in water (cooler temperatures of 29–29.5°C) and maintain cardiovascular fitness (Konlian 1999). Oxygen consumption of an activity performed in water has been shown to be greater than oxygen consumption of the same activity done on land (Cassidy & Nielsen 1992; Fyestone et al. 1993; Routi et al. 1994). Hall et al. (1996) compared water-based exercise, land-based exercise, seated water immersion and progressive relaxation in 139 chronic rheumatoid arthritis patients in a randomized trial. At 3 months, the water-based exercise group maintained the most improvement in emotional and psychological scales. Post anterior cruciate ligament repair patients obtained quicker range of motion gains with water-based exercise when compared with conventional land-based therapy (Norton et al. 1996). It is important to incorporate land-based as well as water-based exercises together in the rehabilitation programme. The goal is to prepare the body for function on land because most people function in life on the ground and do not live their life in the water.
Warm Water Immersion

Heating through the use of submersion in water is a form of convection. Whirlpools are used when a small area of the body is to be heated, such as a part of the upper or lower extremity, while Hubbard tanks are used to treat larger surface areas. The Hubbard tank, due to its larger size, also allows for range of motion manoeuvres. Since larger body areas are exposed to heat during hydrotherapy, there is an increased risk of elevation in core body temperature. Therefore, water temperature rarely exceeds 40°C for total body immersion, whereas temperatures as high as 43°C may be used for partial limb immersion.

The benefits of whirlpool treatment stem from the principle of buoyancy, in which a gravity-eliminated environment assists the patient in upward movement. An additional benefit comes from the resistance to flow, which provides low resistance for muscle strengthening and training. Agitation created by the water flow provides sensory input to the skin, assisting with pain control as well as maintaining appropriate water temperature (Hayes 1984).

As larger areas of the body are immersed in the heated water, diminished regulation of core body temperature occurs, as sweating and heat exchange can only occur in the non-immersed portions (Lehmann & de Lateur 1990). As an area with poor circulation is exposed to heat, a greater demand for blood supply is created due to increased metabolic needs. However, this demand for increased circulation may not be adequately met, leading to ischaemic results. Lowering of water temperature may limit such adverse effects. The indications and contraindications to hydrotherapy are the same as for therapeutic heat.

Paraffin baths

Paraffin baths are used in conjunction with manual joint mobilization, scar desensitization, scar mobilization and active range of motion exercises of the distal upper body or limbs of patients. This may include patients who have hand or foot contractures, healed ankle or hand fractures, postoperative Achilles tendon reconstruction and healed elbow dislocations. Paraffin baths are a mixture of paraffin and mineral oil which provide a useful means of delivering heat to the smaller joints of the body by conduction. The addition of mineral oil creates a lower melting point for the paraffin and therefore provides increased thermal energy release in comparison with water. The bath mixture is kept at a temperature of 52–58°C for upper extremity use and a somewhat lower temperature (45–52°C) for lower extremity application (Griffin & Karselis 1982). Two methods of heating are commonly employed: the dip/wrap method and the immersion method. With dipping, the extremity is immersed in the bath and then withdrawn after a few seconds, thereby allowing the wax to harden. This procedure is then repeated until a glove of wax has been created. Wrapping follows dipping whereby the extremity is enclosed within a plastic or terry towel wrap to create an insulating effect. The glove is then removed after approximately 20 min. The immersion method keeps the extremity within the paraffin bath for a period of 20–30 min. This provides for higher and more prolonged tissue temperatures than that created by the dip and wrap method. Contraindications to paraffin bath use includes patients with open wounds and severe peripheral vascular disease.

Fluidotherapy

Although infrequently used today, fluidotherapy may sometimes be used for postoperative hand and ankle rehabilitation in conjunction with manual joint mobilization, scar desensitization, scar mobilization and active range of motion exercises. Fluidotherapy involves placement of the extremity into a container through which hot air is blown within a medium of dry powder or glass beads (Borell et al. 1977). Treatment provides both therapeutic heat with its associated physiological responses, as well as mechanical stimulation that may further assist with pain control. Borell et al. (1980) found that fluidotherapy, in comparison with hydrotherapy and
paraffin wax, caused the most significant temperature rise in the joint capsule and muscle.

**Diathermy or deep heating modalities**

Diathermy utilizes the principle of conversion to heat deeper tissues. The most commonly used deep heating agents include ultrasound and phonophoresis. Shortwave and microwave diathermy are not used much anymore. Radiant heat is not frequently used any more. The general indications and contraindications are similar to those already discussed for superficial heat. However, specific clinical uses and precautions are presented here.

**Ultrasound**

Ultrasound is defined as sound waves classified within the acoustic spectrum above 20,000 Hz (cycles per second). It is unique among diathermy modalities in that the production of heat is due to a high-frequency alternating electric current (0.8–1.0 MHz) which is converted via a crystal transducer to acoustic vibrations, rather than to electromagnetic energy. Energy transfer occurs due to the piezoelectric effect whereby the crystal undergoes changes in shape when the voltage is applied. By altering the crystal’s configuration, vibrations are created which then pass through the tissues being treated. The heating effects depend on the absorption and reflection of ultrasonic energy based on the differences in the acoustic impedance at the tissue interface. Selective heating is greatest when acoustic impedance is high, such as at the bone–muscle interface. On the other hand, ultrasonic energy is readily conducted through homogeneous structures such as subcutaneous fat or metal implants with minimal thermal effects due to the rapid removal of heat energy. Thus, ultrasound can be safely used in the presence of metal implants. However, in the presence of methyl methacrylate and high-density polyethylene, which may be used in total joint replacements, a greater amount of ultrasound energy will be absorbed with the potential for overheating (Lehmann & de Lateur 1982b). Significant heating can occur to depths up to 5 cm below the skin surface, thereby providing therapeutic effects to bone, joint capsule, tendons, ligaments and scar tissue (Santiesteban 1985). In summary, ultrasound provides for the deepest penetration of all heating modalities since minimal energy is converted to heat in subcutaneous fat or muscle with most of the conversion occurring at the bone interface.

In addition to the thermal effects, ultrasound also has non-thermal effects, which do not relate to tissue temperature elevation but rather to molecular vibration. Although heat can increase membrane permeability, diffusion can also occur due to the non-thermal streaming/stirring effect of fluids created by the ultrasonic field. Gaseous cavitation is also a non-thermal ultrasonic event. Gas bubbles are created as a result of acoustic rarefaction and compression causing subsequent enlargement in bubble size and pressure changes within the tissues. As the gas-filled cavity vibrates due to alternating compression and rarefaction, surrounding fluid movement occurs with the potential for cell destruction. Cavitation can be minimized by the application of external pressure and the use of a stroking, rather than a stationary, technique, which will be discussed shortly.

**Application methods**

Two primary methods of ultrasound application may be used: continuous or stationary. A coupling medium, such as mineral oil/gel or water for irregular surfaces is utilized to ensure adequate transmission of sound energy (Warren *et al.* 1976a; Griffin 1980; Balmaseda *et al.* 1986). With the continuous method, the ultrasound head is moved in a stroking fashion over the area being treated. This provides for safer, more uniform heating. The size of the applicator head should be larger than the treatment field with common sizes ranging from 5 to 10 cm². The stationary method, as the name implies, does not involve the continuous movement of the
applicator head. Since a rapid rise in temperature is produced over a localized area with increased risk of burning and gaseous cavitation, this method is less commonly used. When this method is employed, intensity output is reduced.

Dosimetry is measured in watts per square centimetre (W·cm⁻²), which reflects the applicator output divided by the surface area of the applicator. Intensities of 1–4 W·cm⁻² are most commonly used for the continuous method. Treatment often begins at 0.5 W·cm⁻² and the total output gradually increases. When using the stationary head, a safe range would be 0.1–1.0 W·cm⁻² (Lehmann & de Lateur 1982b). The duration of most treatments is 5–10 min per site based on the size of the treatment area, with 10–12 treatments per series (Hayes 1984). As with all therapeutic modalities, the patient’s subjective response to heating is the best guide for proper dosing.

A variation of the primary technique includes ultrasound application under water to more effectively treat irregular surfaces. Care must be taken to minimize the development of gaseous cavitation over the body part by removing the bubbles formed. Forrest and Rosen (1989) evaluated the effectiveness of heating tendons overlying the lateral epicondyle of a pig by comparing the application of the ultrasound/coupling agent directly over the limb with limb immersion in a water bath with ultrasound application 2 cm from the skin surface. They concluded that the tendon temperature reached the therapeutic range when the ultrasound was used directly over the anatomical area rather than when given under water.

Pulsed application is a method of administering ultrasound waves whereby the energy produced is intermittent. The purpose is to produce the mechanical, non-thermal reactions of ultrasound by allowing for rest periods and subsequent cooling (Prentice 1986). However, evidence is lacking that the non-thermal effects produced by pulsing have any advantage over the similar results produced by the continuous method (Lehmann & de Lateur 1982b).

**MUSCULOSKELETAL CONDITIONS**

General indications and contraindications for ultrasound use in musculoskeletal conditions are the same as for therapeutic heat. Additional precautions include ultrasound over laminectomy sites, the heart, brain, cervical ganglia, tumours, acute haemorrhage sites, pacemakers, infection sites and fluid-filled cavities such as the eyes (Basford 1998). In general, ultrasound may be effective as a therapeutic modality in subacute and chronic inflammation (Hayes 1984). Pain control may occur by both thermal and non-thermal effects. Various studies have shown alteration in nerve conduction velocity after diathermy application, with the changes appearing to be related to energy intensity of the ultrasound field (Madsen & Gersten 1961; Zankel 1966; Currier et al. 1978; Halle et al. 1981). By altering the conduction velocity, analgesic effects may occur.

Studies have also documented increased levels of cortisol following ultrasound application to peripheral nerves. This release may provide increased anti-inflammatory effects (Griffin et al. 1965). However, Gnatz (1989) found that ultrasound applied to the backs of two patients with documented lumbar disc herniation caused increased pain in a radicular pattern. Thus, any pain-relieving effects secondary to cortisol release may be overcome by the increased oedema due to the deep heating effects of ultrasound. Conflicting results have also occurred with the use of ultrasound as a diagnostic technique in the evaluation of lumbosacral nerve root irritation. Cole and Gossman (1980) concluded that ultrasound application over an irritated nerve root provoked pain radiation, providing diagnostically useful results, whereas Reid et al. (1989) concluded that due to the low sensitivity, sonation of the lumbar nerve roots in patients with sciatica was not useful as a screening test for nerve irritation secondary to disc disease.

The literature is filled with uncontrolled, non-randomized studies that show only modest benefit of ultrasound for musculoskeletal disorders.
(Tiidus 1999; Baker et al. 2001; Robertson & Baker 2001). A closer look at just the randomized, controlled studies, however, failed to produce convincing evidence of the efficacy of ultrasound (van der Windt et al. 1999).

Gam et al. (1998) studied the effects of ultrasound, massage and exercise in 58 patients with neck and shoulder myofascial trigger points in a randomized, controlled study. The first group received all three treatments, the second group received massage, exercise and sham ultrasound, and the third group was a control group receiving no treatment. Both of the treatment groups had significantly improved numbers of myofascial trigger points compared with the non-treated control group, but there was no difference between the therapeutic ultrasound and the sham ultrasound groups.

Van der Heijden et al. (1999) studied 180 patients with soft tissue shoulder disorders in a randomized, blinded, controlled trial comparing bipolar interferential electrotherapy with ultrasound as adjuvants to a supervised exercise programme. All 180 subjects received exercise therapy; 73 subjects received active treatments, 72 subjects received dummy ultrasound and interferential electrotherapy, and 35 subjects received no adjuvants. At the 6-week and 12-month follow-ups, the ultrasound and electrotherapy groups had received no additional benefit above the exercise programme alone. Crawford and Snaith (1996) found no difference between ultrasound and sham ultrasound in patients with plantar heel pain. There is conflicting evidence whether ultrasound may be beneficial for delayed-onset muscle soreness (Hasson et al. 1990; Plaskett et al. 1999).

Ebenbichler et al. (1999) studied ultrasound therapy in 54 patients with radiographically confirmed calcific rotator cuff tendinitis in a randomized, double-blinded, sham-controlled study. Thirty-two shoulders were treated with ultrasound and 29 shoulders were treated with sham ultrasound five times a week for the first 3 weeks, then three times a week for the following 3 weeks. At 6 weeks, the treatment group reported greater improvement in pain and quality of life; however, there was no significant difference at the 9-month follow-up. Interestingly, 42% of the ultrasound-treated shoulders demonstrated resolution of calcium deposits and 23% showed improvement. In contrast, the sham group showed calcium deposit resolution in only 8% and improvement in only 12% of subjects (P = 0.002). Perron and Malouin (1997) did not, however, find improvement with ultrasound and acetic acid iontophoresis above the control for their smaller group of 22 patients with calcific shoulder tendinitis. Nykanen (1995) also found no benefit in using ultrasound over sham ultrasound in 72 inpatients with shoulder pain in a randomized, double-blinded, sham-controlled study. Downing and Weinstein (1986) also found no benefit from ultrasound in subacromial bursitis in a double-blinded, sham-controlled study.

Tiidus (1999) reviewed the literature and found little data to support the use of ultrasound for postexercise muscle damage.

There is some evidence to suggest that ultrasound may hasten the healing of bone and tendon. Kristiansen et al. (1997) studied 61 distal radius fractures in a multicentre, prospective, randomized, double-blind clinical trial comparing low intensity non-thermal pulsed ultrasound versus a placebo device. The ultrasound group experienced significantly faster radiographic fracture healing (mean of 61 days in the treatment group vs 98 days in the placebo group). Ramirez et al. (1997) performed work with ultrasound and Achilles tendon injuries of neonatal rats to suggest that ultrasound stimulates collagen synthesis in tendon fibroblasts and stimulates cell division during phases of rapid cell growth. Jackson et al. (1991) also found that ultrasound facilitated the rate of rat Achilles tendon repair by promoting synthesis of collagen which also proved to have a greater breaking strength. Enwemeka (1989) found similar increased tensile strength in the Achilles tendons of rabbits that were treated with ultrasound.

Binder et al. (1986) reported statistically significant results studying 38 patients having ultrasound (1 MHz, 1–2 W·cm⁻²) versus 38 patients
receiving sham ultrasound for lateral epicondylitis. Twenty-four subjects in the treatment group compared with only 11 in the sham group reported satisfactory results. Other studies evaluating ultrasound as treatment for lateral epicondylitis do not show statistically significant improvement (Halle et al. 1986; Lundeberg et al. 1988; Haker & Lundeberg 1991; Vasseljen 1992; Pienimaki et al. 1996).

Gam and Johannsen (1995) and van der Windt et al. (1999) reviewed well-designed trials evaluating ultrasound therapy and concluded that there is little evidence to show efficacy in the treatment of musculoskeletal disorders. Taking the available literature as a whole, there is evidence to support improved tendon and bone healing and only little evidence suggesting a modest benefit of ultrasound in rotator cuff calcific tendinosis and lateral epicondylitis.

**Phonophoresis**

Phonophoresis is sometimes used as a therapeutic modality in the rehabilitation of sports injuries. There are few good controlled studies documenting the clinical effectiveness of phonophoresis. Griffin showed that it is possible to drive a cortisol ointment onto pig skin in situ and to drive it into underlying muscle with ultrasonic energy using levels within the clinical range (Griffin & Touchstone 1963). Newman et al. (1958) reported that hypospray injection of cortisol followed by local application of ultrasonic energy showed an improvement in symptoms of shoulder bursitis compared with ultrasound used alone. Davick et al. (1988) demonstrated that ultrasonically treated topical application of tritiated cortisol can lead to significant increases in cortisol penetration beyond the stratum corneum and into the viable epidermis as compared with topical cortisol alone. They concluded that once beyond the stratum corneum, the cortisol may penetrate over time and be absorbed into the deep soft tissue structures.

In Kleinkort and Wood’s (1975) retrospective study of 285 patients treated for a variety of common inflammatory conditions, phonophoresis using 10% hydrocortisone demonstrated better subjective and objective outcomes as compared with phonophoresis using 1% hydrocortisone.

Klaiman et al. (1998) studied phonophoresis versus ultrasound in 49 patients with various soft tissue injuries in a randomized, double-blinded, uncontrolled trial. Each group had treatments three times a week for 3 weeks; each group had decreased pain levels at the end of 3 weeks, but there was no significant difference between the groups.

In summary, ultrasonic energy in conjunction with topical hydrocortisone may produce some anti-inflammatory effect, possibly related to the use of ultrasound alone and in part as a result of the penetration of cortisone.

**Therapeutic electricity**

The use of therapeutic electricity dates back many centuries. One of the earliest accounts of the use of electricity for a musculoskeletal problem occurred in 1747 when a man with rheumatoid arthritis and involvement of the small joint in his hands received marked relief of his pain symptoms through the use of electricity (Licht 1983). Today, different forms of therapeutic electricity are used in the treatment of musculoskeletal and sports injuries. Electric currents are used to promote healing of injured tissue, to stimulate muscles, to stimulate sensory nerves in treating pain or to create an electrical field on the skin surface to drive ions beneficial to the healing process into or through the skin. This section describes different forms of therapeutic electricity, the scientific basis for their use, their indications and contraindications, briefly describes the techniques used in their applications, and elaborates on their use in sports medicine. Electrical devices can put out either alternating current (AC or faradic), which is the form usually found in household appliances, or direct current (DC or galvanic), which is found in a generator or battery. Direct current can be continuous or intermittent, and can have different waveforms, frequencies, duration and amplitudes. Adjustments in any or all of these parameters will
have an effect on the quality, type and form of stimulation received by the patient. Details of these parameters can be found in the cited texts (Prentice 1986).

It is important to note that the responses muscle and nerve have to electricity vary. Nerve tissue accommodates rapidly to current. Nerve stimulation requires a current which rises rapidly to maximum intensity. High frequencies and short durations are used. Motor nerves respond to 25 cycles·s$^{-1}$ and 500 μs or shorter of stimulus duration. Sensory nerves respond to 100–150 cycles·s$^{-1}$ and 100 μs or shorter stimulus duration. Muscle tissue does not accommodate as rapidly. Muscle can be stimulated with very slowly rising currents. Lower frequency and longer duration stimuli are used in stimulating muscle as compared with nerve. The various forms of therapeutic electricity include TENS, muscle electrical stimulation, percutaneous electrical nerve stimulation, iontophoresis and interferential current.

**Electrical stimulation to promote tissue healing**

Medical galvanism, or the use of galvanic stimulation, uses direct current modalities that deliver a unidirectional, uninterrupted current flow within the tolerance of the patient and without the destruction of tissue. The action of direct current on the body is primarily chemical. This type of modality can be used to directly stimulate muscle following a nerve injury, to produce ionic changes within the tissues and to decrease oedema; or it can be used to introduce topically applied medications into the skin, when it is termed iontophoresis (Hillman & Delforge 1985). The purpose of this electrical stimulation is primarily for the vasomotor effects, i.e. increased circulation. These effects are most often seen in resolution of inflammation, relief of pain and reduction of interstitial oedema through electroosmosis and the shifting of water toward the electrical cathode (Prentice 1986). Motor nerves are not stimulated by a steady flow of direct stimulation. The sensations experienced result in part from accumulation of ions in the skin under the electrodes, which act as a physiological stimulus to the sensory nerve endings, producing reflex vasodilatation (Marino 1986).

Direct, uninterrupted electric current tends to be quite uncomfortable and may cause superficial skin burns. For this reason, a modification of the technique has been developed whereby the current is applied in an alternating or ‘pulsed’ manner, termed high-voltage pulsed galvanic stimulation (HVPGS). Although the main use of HVPGS in sports rehabilitation is for relief of pain, it can also be used to aid in tissue healing (Ross & Segal 1981). Electrically induced muscle contraction, such as that obtained with direct current or HVPGS, can be used to duplicate regular muscle contractions. These contractions help stimulate circulation by pumping blood through venous and lymphatic channels after acute injuries, when fluid accumulation is significant. Intermittent muscle contraction, which permits increased blood flow, may also produce relaxation of the muscles. Electrical stimulation of muscle contractions can allow resolution of inflammatory fluid, while keeping an injured joint protected. In order to be successful in reducing swelling, the current intensity must be high enough to provide a strong, comfortable muscle contraction, therefore interrupted or surge-type pulses must be used (Prentice 1986).

There are a number of contraindications to the use of electrical current in sports rehabilitation. Contraindications of electrical therapy include stimulation over cardiac pacemakers, electrical implants, carotid sinus, epiglottis, abdomen and gravid uterus (Basford 1998). Treatment should be avoided over any area that is anaesthetic to avoid local burns. Recent scars in the area to be treated should also be avoided because of the potential for wound dehiscence. Any area where metal is embedded close to the skin in the area to be treated should be avoided for fear of concentrating the heat source at the metal surface. Any form of electricity should be avoided near an area of acute injury if active bleeding is still present to prevent worsening of the haemorrhage.
Iontophoresis: driving ions with electrical stimulation

Iontophoresis uses direct current to drive medicinal ions locally into the skin and mucous membranes. Ions of the medicinal compounds are absorbed subcutaneously. This absorption occurs slowly into the local soft tissue, while some is ultimately absorbed systemically. Evidence for penetration much beyond the skin is variable and may depend on the particular substance (O’Malley & Oester 1955). James et al. (1986) showed that percutaneous iontophoresis of 1% prednisolone sodium phosphate through human skin and nails gives peak plasma levels of about one-third that produced by oral ingestion of 10 mg prednisolone. Similarly, Zankel et al. (1959) showed that the absorption of the negative ion I through unbroken skin only occurred in those conditions which included iontophoresis. Bertolucci (1982), in a double-blind study, found that patients below the age of 45 years with shoulder dysfunction related to primary tendonitis responded well to iontophoresis steroid administration, whereas the placebo group received no benefit from the iontophoretic treatment with sodium chloride. However, Chantraine et al. (1986), in studies done both in vivo and in vitro failed to demonstrate the transcutaneous migration of corticosteroids with iontophoresis.

Iontophoresis has been used to drive multiple substances into the skin, some of which include calcium chloride, hydrocortisone, lithium chloride, lidocaine and acetic acid. Acetic acid theoretically replaces carbonate in calcium carbonate to become calcium acetate, which is blood soluble, thereby leaching the calcium away from the site of bony spur and inflammation. Japour et al. (1999) reported very encouraging results using acetic acid iontophoresis in 35 patients with chronic heel pain in an uncontrolled case series. Ninety-four per cent of patients reported relief of heel pain, after an average of only 5.7 sessions, over an average 2.8-week treatment period. Pain scores decreased from 7.5 to 1.8, which continued to be low at 0.64 at the 27-month follow-up. Without a control group, however, it is not known whether these encouraging results are from treatment effect or the natural course of heel pain.

Perron and Malouin (1997) did not, however, find improvement with ultrasound and acetic acid iontophoresis above the control in their small group of 22 patients with calcific shoulder tendinitis. Gudeman et al. (1997) studied 0.4% dexamethasone iontophoresis in the treatment of 40 feet with plantar fasciitis in a randomized, double-blinded, placebo-controlled trial. After six treatments over 2 weeks, the treatment group significantly improved, but were no different than placebo at the 1-month follow-up. Taniguchi et al. (1995) reported a significantly increased pain threshold in 30 healthy volunteers with clonidine iontophoresis, as compared with amitriptyline and imipramine iontophoresis; the clinical implications of this finding are not known. Hasson et al. (1992) studied 18 females with delayed-onset muscle soreness with dexamethasone iontophoresis, placebo iontophoresis and no treatment. Perceived muscle soreness was significantly less in the treatment group 48 h later, but no change in strength was noted.

The clinical effectiveness of iontophoresis is still debatable. Some of the claimed benefits for pain relief may be due to the effects of the direct current used as opposed to the medicinal compounds purportedly driven into the circulation. Few side effects have been described other than drug sensitivity and sensitive skin.

Electrical stimulation of muscle

Electrical stimulation of muscle is accomplished with either direct or alternating current, or a combination of the two. Alternating current is usually preferred to direct current because of greater patient comfort (Hillman & Delforge 1985). Muscles are stimulated for one of four reasons. The first is to aid in muscle pumping for oedema reduction and tissue healing as described in the previous section. The other reasons include re-educating muscle, retarding atrophy in immobilized or partially denervated muscle and enhancing strength.
Certain general principles need to be adhered to when performing electrical stimulation of muscle (Stillwell 1982). Good contact should be maintained between the skin and the electrodes. The active electrode should be placed over the motor point of the muscle. The two electrodes used should generally be placed on the same side of the body. Finally, since denervated muscle does not have a motor point, the active electrode may be placed at the point that gives the best motor response, or the two electrodes may be placed one at each end of the muscle so that the current will pass through the muscle and stimulate all of it.

**Muscle Re-Education**

Electrical muscle stimulation can be used for muscular re-education after sports injuries (Amrein et al. 1971). Muscular inhibition is quite common after traumatic injuries or surgery. Central nervous system inhibition is often the cause, as muscle contraction causes pain. This then causes decreased contraction and ultimate immobilization of the affected muscle. The injured patient or athlete may have a difficult time initiating contraction of an injured muscle because of the pain associated with movement and the lack of sensory input from that muscle due to disuse. Forcing the muscle to contract causes an increase in the sensory input from the muscle, allowing the patient to see the muscle contract, and then attempt to duplicate this muscular response (Prentice 1986). The focus of this type of training is on kinetic training and the sensory awareness of muscular contraction. For muscle re-education to occur, the current intensity must be adequate for a muscle contraction, but not too uncomfortable for the athlete. HVPGS or high-frequency alternating current may be most effective (Prentice 1986). Although the clinical implications are not clear, electrical muscle stimulation has been shown to selectively increase strength when applied to the abdominal muscles (Alon et al. 1987), triceps brachii (Snyder-Mackler et al. 1988) and erector spinae (Kahanovitz et al. 1987). Improved motor performance was demonstrated in the deltoid and pectoralis major (Fleury & Lagasse 1979) and in the abductor hallucis (LeDoux & Quinones 1981).

**Retardation of Muscle Atrophy**

Electrical muscle stimulation may help prevent strength losses (Eriksson & Haggmark 1979; Godfrey et al. 1979; Gould et al. 1979; Grove-Lamine et al. 1983; Morrissey et al. 1985; Wigerstad-Lossing et al. 1988; Delitto et al. 1989; Snyder-Mackler 1990; Abdel-Moty et al. 1994), as well as prevent muscular atrophy (Eriksson & Haggmark 1979; Morrissey et al. 1985; Nitz & Dobner 1987; Wigerstad-Lossing et al. 1988), which occur when a limb is immobilized. Stanish et al. (1982) and Eriksson et al. (1981) have shown that the biomechanical changes occurring in the muscles of immobilized limbs are retarded by electrical muscle stimulation. Eriksson and Haggmark (1979) showed better muscle function in a group of patients after reconstruction of the anterior cruciate ligament when treated with electrical muscle stimulation and isometric exercise than with isometric exercise alone. However, Halbach and Straus (1980) examined isokinetic only and electrical stimulation only exercise programmes and found that the isokinetic workload was greater than the electrical stimulation workload after 3 weeks of training and stimulation.

In general, it is agreed that electrical muscle stimulation programmes are more effective than no exercise programme, but no more effective than traditional strengthening exercise programmes (Currier et al. 1979; Halbach & Straus 1980; Kramer & Mendryk 1982; Currier & Mann 1983; Kramer & Semple 1983; Laughman et al. 1983; McMiken et al. 1983; Kubiak et al. 1987; Lieber et al. 1996). Snyder-Mackler et al. (1995) showed there may be some functional benefit from using high-intensity muscle stimulation in addition to an exercise programme. They studied 110 patients following anterior cruciate ligament reconstruction and randomized them to receive high-intensity neuromuscular electrical stimulation, high-level volitional exercise, low-intensity neuromuscular electrical stimulation, or combined high- and low-intensity neuromuscular
stimulation in addition to a standard closed kinetic chain strengthening programme. At 4 weeks, the high-intensity stimulation group enjoyed the best quadriceps strength gains of 70% of the uninvolved side, as opposed to 57% in the high-level volitional exercise group. With knee joint kinematic analysis, the high-intensity stimulation group showed significantly more normal knee flexion–extension sagittal plane excursions during the stance phase when compared with the other groups (Fig. 10.2). However, long-term benefits and other functional outcomes were not studied.

In order to improve strength under any circumstances, either electrically stimulating muscle or through voluntary contraction, maximal or near maximal contractions must occur to the point of muscle fatigue. The discomfort of the stimulation remains a major limitation (Currier 1991). Most research to date indicates that, with few exceptions, maximal contractile forces can be produced by voluntary contractions (Singer et al. 1989).

Also, while electrical stimulation retards denervation atrophy, its effect depends on the pulse duration, the frequency and intensity of the current, the placement of the stimulating electrode, the duration and number of treatment sessions, the rest periods between the treatment sessions and the resting length of the muscle during stimulation. Further discussions of all these parameters are beyond the scope of this chapter, but are well described by Licht (1983). Gibson et al. (1988) suggest that electrical stimulation seems to prevent the fall in muscle protein synthesis that is related to immobility. Cabric et al. (1987) found that lower frequency electrical stimulation (50 Hz) of muscle increased muscle fibre size, which was thought to be correlated to the proliferation of muscle cell nuclei.

**Muscle Strength Gains**

Increasing the strength of a muscle can be accomplished by repeated maximal or submaximal contractions of that muscle. The rationale behind the strength enhancement is that electrical stimulation can either increase the maximum contractile force in the muscle, or that it can recruit more fibres to contract with a given stimulus, thereby enhancing the strength of contraction (Singer et al. 1989). When done via electrical stimulation, tetanic contractions, which are achieved at pulse rates above 20–30 per second, are required for maximum muscle contraction. Electrical stimulation is achieved by stimulating the motor nerve to a muscle by means of electrodes placed on the skin. Most of the research done with muscle stimulation for strength gains has been done with electrical stimulation in the isometric mode. Because of the many variables of stimulation of muscle—including frequency, duration, pulse shape, intensity and charge—no studies exist that compare the effectiveness of these variables in inducing effective strength-improving muscle contractions (Singer et al. 1989).

Significant strength gains in normal muscles have been described by Kots (1977). Kots’ ‘Russian’ stimulation used an alternating-type current of high pulse rate and high intensity to produce strong, involuntary muscle contractions with associated stimulation of local blood flow (Hillman & Delforge 1985). The actual benefit of such a
programme is questionable as similar results have not been duplicated. To date, there are no good studies to warrant electrical stimulation for gaining strength in normal muscle.

There has also been some suggestion that neuromuscular electrical stimulation can help in controlling oedema after injury (Gould et al. 1979; Lake 1989; Griffin et al. 1990). Hsueh et al. (1997) studied 60 patients with upper trapezius myofascial trigger points in a randomized, placebo-controlled trial. They found electrical muscle stimulation significantly improved immediate cervical range of motion over the placebo and electrical nerve stimulation groups. The electrical nerve stimulation group, however, showed significantly decreased immediate pain scores when compared with the other two groups. Pope et al. (1994) showed no statistical difference in physical measures in a randomized trial of 164 patients with subacute low back pain who received 3 weeks of electrical muscle stimulation, manual manipulation, massage or corset wearing.

In summary, the literature suggests that electrical muscle stimulation may be helpful in strengthening normal muscle, preventing loss of muscle bulk and strength associated with immobilization, selective strengthening, enhancing motor control and controlling oedema after injury (Lake 1992).

**Electrical stimulation of nerves**

The use of alternating and direct current for pain reduction via nerve stimulation is commonly used in the rehabilitation of sports injuries. The goal of nerve stimulation is to stimulate sensory nerves to change the patient’s perception of a painful stimulus coming from an injured area. Nerves, being more sensitive to electric current than muscle, are usually stimulated at a high frequency and short duration. Three theories have been put forward to explain the analgesic effects of electrical stimulation—the gate-control theory, the central biasing theory and the opiate pain control theory.

Melzack and Wall (1965) proposed the gate-control theory in 1965. It postulates that when large-diameter afferent nerve fibres are electrically stimulated, painful stimuli arriving at the spinal cord through small fibres at the dorsal horn are blocked from transmission to the central nervous system where pain is perceived. By stimulating the large fibres, the ‘gate’ to allowing further impulses to be transmitted centrally is closed.

The central biasing theory also uses the idea of gating impulses. In this model, intense stimulation of smaller fibres (C fibres or pain fibres) at peripheral sites, for short periods, causes stimulation of the descending neurones, which then affects transmission of pain information by closing the gate at the spinal cord level (Prentice 1986).

The opiate pain control theory is premised on the fact that endogenous opiate enkephalins and β-endorphins exist in our systems. Electrical stimulation of sensory nerves may stimulate the release of these compounds from local sites throughout the central nervous system, causing pain relief.

**TRANSCUTANEOUS ELECTRICAL NERVE STIMULATION (TENS)**

TENS is usually defined as the application of an electric current therapy through the skin to a peripheral nerve or nerves for the control of pain (Singer et al. 1989). TENS units are very commonly used in sports rehabilitation (Roesser et al. 1976). There are a number of advantages of using TENS, including that it is comfortable, fast-acting, can be used in many types of pain problems, and can be used continuously. The disadvantages are that the carry-over is often variable and adaptations may need to be made for continued benefit, i.e. increasing the pulse width or amplitude.

There are two major forms of TENS—high frequency and low frequency. High-frequency TENS stimulates sensory nerves and causes an increase in pain threshold, although it does not stimulate the release of endorphins. Low-frequency TENS will cause pain relief, which can be blocked by administration of opiate antagonists, implicating the opiate theory of pain...
control. It is sometimes likened to acupuncture. Walsh et al. (1995) found that low-frequency TENS (4 Hz) was the best in decreasing immediate pain visual analogue scores for ischaemically induced pain in a randomized, double-blinded, controlled trial of 32 subjects comparing 4 Hz TENS, 110 Hz TENS, placebo and no treatment. When applying electrodes for either type of TENS, electrodes may be placed on or around the painful area, over specific dermatomes, myotomes or sclerotomes that correspond to the painful area, close to the spinal cord segment that innervates an area that is painful, or over trigger or acupuncture points.

The efficacy of TENS is not clear-cut (Licht 1983; Robinson 1996). Although it has been argued that blinding for TENS studies may be difficult to truly achieve (Deyo et al. 1990b), few studies have been done in a double-blind manner (Gersh 1978; Robinson 1996). Thorsteinsson et al. (1977) studied 93 patients with various disorders in a double-blind, randomized, sham-controlled, cross-over study. Forty-nine per cent of the TENS group versus 32% of the sham group reported partial or complete pain relief, but this difference was not statistically significant.

Ordog (1987) studied 100 patients with acute traumatic disorders—including sprains, fractures and contusions—in a randomized study comparing four groups: active TENS plus acetaminophen/codeine, sham TENS plus acetaminophen/codeine, active TENS and sham TENS. Pain levels in 2 days dropped 63% in the active TENS plus drug, 58% in the sham TENS plus drug, 45% in the active TENS, and 17% in the sham TENS groups. There was no statistical difference between the active TENS and the sham TENS plus drug groups suggesting that TENS may be as effective as acetaminophen/codeine for acute traumatic disorders.

Lehmann et al. (1986) studied 54 patients with chronic low back pain in a randomized study of active subthreshold TENS, sham TENS or electroacupuncture. They found that although all groups improved from baseline, there were no statistically significant differences between the groups at 6 months.

Deyo et al. (1990a) studied 145 volunteers with chronic low back pain in a randomized trial comparing four groups: active TENS, sham TENS, active TENS with exercise and sham TENS with exercise. At 8 weeks following the 4-week treatment period, there were no significant differences between the active TENS groups and the sham TENS groups.

Marchand et al. (1993) studied 43 patients with chronic low back pain in a randomized sham TENS (placebo) and no intervention (‘nocebo’) controlled study. Immediately after two treatments of 30 min each, the active TENS group showed a 43% reduction of pain intensity as compared with a 17% reduction in the sham TENS group—this finding was statistically significant. Active TENS patients maintained significantly decreased pain intensity scores at 1 week as compared with sham TENS patients. Both sham and active TENS groups’ pain intensity scores were significantly better than the no treatment group at the 3- and 6-month follow-ups, but there was no statistically significant difference between the sham TENS and active TENS groups.

Herman et al. (1994) studied 58 work-related acute low back injuries of 3–10 weeks duration in a randomized study of active TENS versus sham TENS. Both groups received the same exercise programme. Outcome measures showed the active TENS group produced a drop in visual analogue pain scores immediately after TENS but showed no difference in disability, return to work rates or pain scores at the 4-week follow-up.

Jensen et al. (1985) showed that a group of patients treated with TENS following arthroscopic surgery experienced less pain, required less narcotics and regained strength more quickly than the control group. Sluka et al. (1998) showed that high-intensity TENS was more effective in providing immediate pain relief than low-intensity TENS for induced knee pain following injection.

Some studies have shown that TENS may provide short-term relief for delayed-onset muscle soreness (Denegar & Huff 1988; Denegar et al. 1989; Denegar & Perrin 1992), but others have not (Craig et al. 1996). TENS can also be helpful for
postoperative pain (Smith et al. 1983; Cornell et al. 1984; Jensen et al. 1985; Arvidsson & Eriksson 1986; Carrol & Badura 2001).

Overall, clinical experiences seem to show that TENS appears to give the best benefit in the treatment of early postoperative pain or pain from acute injuries; however, the response can be variable, unpredictable and short lived. There are no good studies to prove the benefits of TENS in chronic pain problems. It appears that TENS is not harmful, may sometimes provide pain relief above the placebo response, is more effective in acute pain relief with short-term benefit, and is more effective at higher stimulation intensities (Robinson 1996).

No serious complications have been observed from the use of TENS (Licht 1983). Hypersensitivity of the skin has been observed in up to 10% of patients, either due to the electrode or the electrode jelly. Minor burns were observed when high-frequency stimulation and high-intensity stimulation were used at the same time. TENS should not be used in patients with pacemakers, stimulation should not be applied over the carotid sinus, and it should probably not be used during pregnancy, as the effects are unknown (Singer et al. 1989).

**INTERFERENTIAL CURRENT**

Interferential current is another type of electrical stimulation that is used to control pain. Electric signals from two sets of electrodes with the same waveform are applied so that they arrive at the point to be stimulated from two directions. The area where the current overlaps is called the interference pattern, and the intensity summates in a manner that can be effective in modifying pain (Singer et al. 1989). The frequency obtained at the interference pattern, through cancellation of some waveforms, allows stimulation of local muscle and nerve at a greater depth than if applied only from the surface at that point. Thus, interferential current uses high-frequency carrier circuits to afford deeper penetration. The main use of interferential currents in sports is for pain reduction.

**PERCUTANEOUS ELECTRICAL NERVE STIMULATION (PENS)**

PENS combines the therapeutic effects of TENS and electroacupuncture by placing small-gauge needles (32 gauge) 2–4 cm into soft tissues and muscles in a dermatomal distribution, with subsequent stimulation with electrical impulses (< 25 mA) at a frequency of around 4 Hz. Ghoname et al. (1999b) investigated the efficacy of PENS. They studied 60 patients with stable chronic non-radicular low back pain from radiographically confirmed degenerative disc disease in a randomized, single-blinded, sham-controlled, cross-over study. Each subject received all of the following four treatments in a randomized order: sham PENS, PENS, TENS and exercise. Sham PENS consisted of needle placement but no electrical stimulation. Exercise only consisted of one exercise where the patient performed spinal flexion and extension while sitting in a chair. At 24 h, the PENS group showed significant improvements on pain, activity, sleep, sense of well being and narcotic usage scores as compared with the other three groups. Ninety-one per cent identified the PENS treatment as the preferred pain therapy and 81% identified it as a treatment for which they would be willing to pay extra. Ghoname et al. (1999a) also found that a stimulus frequency of 15–30 Hz is more effective than 4 or 100 Hz.

There are numerous potential uses of therapeutic electricity in the rehabilitation of sports injuries. The most common indications are for stimulation of muscle to retard atrophy, to aid in tissue healing and muscle re-education, the stimulation of nerves for pain modification, and to drive ions through the skin for their medicinal effects (Table 10.2).

**Traction**

Traction is the technique in which a distractive force is applied to a part of the body to stretch soft tissues and to separate joint surfaces or bone fragments (Hinterbucher 1985). Cyriax popularized traction in the 1950s as a treatment for
lumbar disc lesions. Prior to that, traction was mainly used in the treatment of fractures. Over the years, traction has gained some popularity in the field of sports rehabilitation, particularly with respect to cervical and lumbosacral spine injuries.

Cervical traction

Cervical traction is used for a number of types of injuries of the cervical spine including cervical herniated nucleus pulposus, radiculopathy, strains, zygapophyseal joint syndromes and myofascial pain. The main reason for its use is relief of pain. Pain relief may occur through one of several mechanisms, including rest through immobilization and support of the head, distraction of the zygapophyseal joints and associated improved nutrition to the articular cartilage, tightening of the longitudinal ligament and decreasing the intradiscal pressure. The last two both press a bulging disc more centrally, relieving nerve root pressure via increased foraminal diameter, improving head posture, and elongating muscles to improve blood flow and reduce spasm.

Determination as to whether the patient has any contraindications is the first step in administering traction of the cervical spine. Contraindications to the use of spinal traction include an unstable spine, ligamentous instability, verteobasilar artery insufficiency, atlantoaxial instability, rheumatoid arthritis, osteomyelitis, discitis, neoplasm, severe osteoporosis, untreated hypertension, severe anxiety, cauda equina syndrome and myelopathy (Rechtein et al. 1998). Clinically, the examiner can first try manual traction with the patient for a few minutes to see what kind of therapeutic response they can expect with traction.

To achieve separation of the spinal segments, a force of significant magnitude and duration must be exerted. Traction can be delivered manually, through weights and pulleys or via a mechanical device. The direction of pull can be vertical, horizontal or at an angle. Traction can be performed while the patient is standing, sitting, lying on a horizontal or inclined plane, or in the prone or supine position. Traction can be continuous, sustained, intermittent or intermittent/pulsed. Surface resistance to traction is dependent on the weight of the body or body segment undergoing traction and the size, quality, contour and texture of the two surfaces in contact (Hinterbucher 1985).

With cervical traction, the optimal angle of pull to obtain the most distractive force with the least weight is at 20–30° of head flexion (Fig. 10.3). The supine position may be more effective than sitting, as it allows more relaxation. In the supine position the force must be sufficient enough to overcome friction and must have a pull that is at least half the weight of the head; 3.6–4.5 kg is a usual starting point. In sitting, the pull must be sufficient to support the head; usually 11–18 kg is necessary.

Cervical traction can be continuous or intermittent. Continuous traction will allow quieting of the stretch reflex and decrease muscle guarding. It will also allow separation of the posterior structures (zygapophyseal joints) if maintained for at least 7 s at a time. Intermittent traction is believed to act by cyclically causing muscle contraction and relaxation, thereby increasing blood flow in a ‘massage-like’ action.

---

Table 10.2 Types of therapeutic electricity and typical uses.

<table>
<thead>
<tr>
<th>Type of Therapeutic Electricity</th>
<th>Typical Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating current or faradic stimulation</td>
<td>Muscle re-education</td>
</tr>
<tr>
<td>Direct current or galvanic stimulation</td>
<td>Retard muscle atrophy</td>
</tr>
<tr>
<td>High-voltage pulsed galvanic stimulation</td>
<td>Improve muscle strength</td>
</tr>
<tr>
<td>Iontophoresis</td>
<td>Pain control</td>
</tr>
<tr>
<td>Transcutaneous electrical nerve stimulation</td>
<td>Pain control</td>
</tr>
<tr>
<td>Percutaneous electrical nerve stimulation</td>
<td>Pain control</td>
</tr>
<tr>
<td>Interferential current</td>
<td>Pain control</td>
</tr>
</tbody>
</table>
Goldie and Landquist (1970) carried out a blind study of patients with cervical pain, using traction, exercise and control groups. Results showed no difference between the three groups, although there were a somewhat larger number of improved patients among those who were subjected to traction. Zylbergold and Piper (1985) studied 100 patients with disorders of the cervical spine in a randomized, controlled trial comparing static traction, intermittent traction, manual traction and no traction. At 6 weeks, all groups improved significantly compared with baseline, but the intermittent traction group scored significantly better than the no traction group in pain and cervical range of motion measures. Swezey et al. (1999) showed that 81% of their 58 patients with cervical spondylosis had symptomatic relief with home cervical traction in a retrospective, uncontrolled study.

Lumbar traction

The traction load necessary to produce vertebral separation in the lumbar spine is much greater than that required to produce vertebral separation in the cervical spine. De Seze and Levernieux (1951) estimated that a tractive force of 330 kg was required to obtain a separation of 1.5 mm at the L4/5 vertebral level and that 365 kg were required to obtain a separation of 2 mm at the L3/4 level. Numerous other studies discuss the great amount of force necessary to obtain separation of even a small amount in the lumbar spine (Frazer 1954; Lawson & Godfrey 1958). Therefore, clinically, lumbar traction functions to limit the activity of the patient. It may assist to decrease muscle spasm by forcing rest. There is no evidence that lumbar traction facilitates nuclear migration of the disc.

Lumbar traction is set up in a similar manner to cervical traction except that a corset rather than a head halter is used. Various types of lumbar traction are described including gravity inversion traction, gravity lumbar reduction and autotraction. Inversion traction must be avoided in patients with hypertension or retinal problems because of the potential for increasing blood and retinal pressures. Generally, at least one-quarter of the bodyweight must be used just to overcome the friction of lumbar traction. The maximum force that a patient can tolerate is often used (Hinterbucher 1985).

Although ample experimental evidence that a traction force of sufficient magnitude and duration applied to the spine produces separation of the vertebrae and zygapophyseal joints and increases the size of the foramen, no clear scientific evidence of its therapeutic value has been reported in the few controlled studies to date (Hinterbucher 1985). Christie’s (1955) controlled study of traction in the treatment of acute and chronic lumbar pain—with and without root

Fig. 10.3 Cervical traction mechanical unit.
signs—showed that traction, when effective, was most useful in patients with chronic backaches with root signs. Weber’s (1973) double-blinded, controlled study of patients with sciatica from a prolapsed disc treated with traction or simulated traction failed to show any significant difference in pain, mobility of the lumbar spine or the presence of neurological signs in either group. Beurskens et al. (1997) studied 151 patients with non-specific low back pain in a randomized, sham-controlled trial which showed no difference between the treatment and control groups. Van der Heijden et al. (1995) reviewed only randomized clinical trials studying traction. Beneficial effects of lumbar traction were suggested in only three of 14 randomized studies and in only one of the 11 studies that the authors considered to be of better quality. They concluded that there was no evidence to support or refute the use of traction.

Although hard data supporting the use of traction are not available, many clinicians and patients will speak of its benefits. Some of these positive experiences may be related to enforced bed rest or the effects of some active intervention for their back complaints. Cheatle and Esterhai (1990) surveyed 369 orthopaedic surgeons and 165 physiatrists about their use of pelvic traction. Twenty-eight per cent of the responding physicians said they would prescribe traction; the chief rationale (54% of respondents) of this prescription was to ensure bed rest. Similarly, there are a lack of scientific data that traction is not physiologically sound or even harmful.

Magnetic therapy

Magnetic therapy has been used for pain relief for centuries and is replete with anecdotal successes. The strongest scientific evidence in support of the therapeutic use of magnetic fields in treating musculoskeletal disorders involves using pulsating magnetic fields of strengths over 10 G. The World Health Organization reported that there is no available evidence to show adverse effects on humans exposed to static magnetic fields up to 2 T, which is equal to 20 000 G (United Nations Environment Programme 1987). The mechanism of action of possible therapeutic effects is unknown (Blank & Findl 1987; Ayr-apetyan et al. 1994). Increased peripheral blood flow has been proposed as a mechanism of action by Erdman (1960), Fenn (1969) and Ross (1990). The increased blood flow may also be associated with changes in fibroblast concentration, fibrin fibres and collagen at wound sites (Goldin et al. 1981). Other postulates include blockage of action potentials (McLean et al. 1995), the role of water (Carpenter & Ayrapetyan 1994), resonance, ions, induction, subatomic magnetic field interactions, and closed electrical circuits within endothelium (Hazlewood & VanZandt 1995).

Pulsating electromagnetic field therapy has shown some promise in the treatment of osteoarthritis of the knee and spine (Miner & Markoll 1993; Trock et al. 1994). Trock et al. (1994) studied 86 subjects with knee osteoarthritis and 81 subjects with cervical spondylosis in a randomized, double-blinded, placebo-controlled trial. The group treated with pulsating electromagnetic fields (10–20 G) showed significant changes from baseline at a follow-up interval of 1 month; this change was not seen in the placebo group. Foley-Nolan et al. (1992) studied the use of pulsating electromagnetic therapy in acute whiplash syndrome in a double-blinded, randomized, controlled study. Twenty patients wore active pulsed electromagnetic therapy collars and 20 patients wore placebo collars for 8 h·day−1. The treatment group had statistically significant improved visual analogue scales at 2 and 4 weeks compared with placebo. Pujol et al. (1998) randomized a small group of 30 patients to receive 40 min of real or sham current magnetic coil stimulation over localized tender body regions. Postprocedure pain scores decreased by 59% in the treatment group versus only 14% in the sham group—this finding reached statistical significance.

Non-union fracture rate healing can be improved with pulsating high magnetic fields (O’Connor et al. 1990; Bassett 1994). Sharrard (1990) studied electromagnetic therapy in the treatment of delayed union tibial shaft fractures using active electromagnetic stimulation units.
in 20 patients versus dummy control units in 25 patients. Nine subjects in the treatment group and only three subjects in the control group showed union of the fracture, which was a statistically significant finding.

There are fewer studies to support the use of static magnets placed over the skin of the affected areas. Washnis (1998) cites a multitude of studies using static magnets, but few have been randomized, double-blinded, placebo-controlled trials appearing in peer-reviewed journals (Vallbona & Richards 1999). No controlled studies have been conducted in an athletic population.

Conclusions

1 Therapeutic modalities play a small, but important, part in the rehabilitation of sports injuries. Knowledge of their indications, contraindications, methods of application and efficacy is essential if they are used.

2 Cold therapies have many applications in sports rehabilitation. Cryotherapy can be used alone in any of its application methods or in conjunction with heat in contrast baths.

3 Heat modalities work via conduction, convection or conversion. Heating can be accomplished through superficial heat modalities, such as hydrocollator packs, hydrotherapy, radiant heat, paraffin baths and fluidotherapy. Heating can also be accomplished through deeper heating modalities, such as ultrasound, phonophoresis and shortwave and microwave diathermy.

4 Therapeutic electricity can be used to promote tissue healing, to stimulate muscles, to stimulate nerves for pain reduction or to drive medicinal compounds into circulation. Stimulation of muscle fibres can be used for muscle re-education, to retard muscle atrophy or possibly to gain muscle strength. Stimulation of nerve tissue is accomplished through the use of TENS, interferential current or PENS.

5 Traction is occasionally used in sports rehabilitation, particularly with neck and low back injuries. The most useful indication for traction is with cervical injuries. The efficacy of lumbar traction is not well proven.

6 Pulsating high magnetic fields may be helpful in osteoarthritis and bone healing, but there are few data to support the use of static magnets for musculoskeletal disorders.

7 Modalities may be used as an adjunct to the rehabilitation of musculoskeletal injuries. There are no definite scientific data to strongly support their use, certainly not as a primary treatment option. They may, however, help to decrease pain and oedema to allow an exercise-based rehabilitation programme to proceed.

References


Deyo, R., Walsh, N., Martin, D., Schoenfeld, L. & Ramamurthy, S. (1990a) A controlled trial of
transcutaneous electrical nerve stimulation (TENS) and exercise for chronic low back pain. New England Journal of Medicine 322 (23), 1627–1634.


Introduction

Musculoskeletal injuries form a large proportion of the injuries reported by athletes (Wang et al. 1993). There are many potential causes for these injuries and these include muscle strength imbalances, muscle fatigue, biomechanical abnormalities, incorrect training methods and musculoskeletal inflexibility (Burkett 1970; Burry 1975; Cooper & Fair 1978; Zarins 1982; Agre 1985). A decrease in musculoskeletal flexibility has been associated with both the aetiology of these injuries, as well as a consequence of these injuries (Nicholas 1970; Glick 1980; Toft et al. 1989; Wang et al. 1993). Furthermore, it has been documented that flexibility training (stretching), using a variety of techniques, effectively increases musculotendinous unit (MTU) range of motion in human subjects (De Vries 1962; Tanigawa 1972; Medeiros et al. 1977; Henricson et al. 1984).

Because there are established guidelines as to the norms for musculotendinous unit flexibility for most joints or groups of joints (Janda 1983; Kendall & McCreary 1983; Travell & Simons 1983; Clendaniel et al. 1984), range of motion testing can reveal values that exceed the normal (hypermobility), or are less than the normal (hypomobility or inflexibility). Hypermobility will not be discussed in detail in this chapter. However, the causes for hypermobility are largely genetic. Also, there are instances where hypermobility is related to certain athletic activities, such as gymnastics or ballet dancing (Grahame & Jenkins 1972). In addition, in general, larger ranges of motion are present in females compared with males, and in younger compared with older individuals (Nicholas 1970). Other causes for hypermobility are an increased musculotendinous unit temperature (Wessling et al. 1987; Anderson & Burke 1991) and, in some instances, after injury or surgery, significant lengthening of soft tissue structures can also result in abnormally increased joint range of motion.

The use of stretching techniques as a means of treating hypomobility is widely advocated, as a means of preventing injuries, as well as forming an important component in the design of a rehabilitation programme following injury or surgery (Cureton 1941; Weber & Kraus 1949; De Vries 1962; Nicholas 1970; Cahill & Griffith 1978; Cooper & Fair 1978; Glick 1980; Beaulieu 1981; Wiktorsson-Moller et al. 1983; Hardy 1985; Etnyre & Abraham 1986; Williford et al. 1986; Toft et al. 1989; Zachazewski 1990; Worrell et al. 1991, 1994). A number of health professionals are therefore involved in the prescription of flexibility training in the primary and secondary prevention of injuries in athletes. These professionals include sports physicians, physiatrists (rehabilitation medicine practitioners), orthopaedists, physical therapists, rehabilitation workers, physical educators, athletic trainers, coaches and athletes.

This chapter focuses on the role of flexibility training in the rehabilitation of injuries. After a brief review of definitions and terminology pertaining to flexibility training, the general benefits of flexibility training and the pathophysiology of loss of range of motion after injury will
be briefly reviewed. Techniques for the measurement of flexibility, and the physiology of stretching will be discussed. A scientific basis for stretching will be presented, and practical guidelines for the implementation of a flexibility training programme during rehabilitation will be provided.

**Terminology and definitions**

The literature uses a variety of terms when flexibility training is discussed. It is therefore necessary to define these terms at the beginning of this chapter in order to clarify what will be referred to in the text that follows. A summary of the terminology and definitions used is depicted in Table 11.1.

The terms ‘musculotendinous unit length’, ‘joint range of motion’ and ‘musculotendinous unit flexibility’ are often used interchangeably. **Flexibility** can be defined as the range of motion available in a joint or in a group of joints that is influenced by muscles, tendons and bones (Anderson & Burke 1991). Flexibility has also been described as the degree to which muscle length permits movement over that which it has an influence (Toppenburg & Bullock 1986).

Measurement of joint range of motion is thereby an indirect measure of the extensibility of the tissues that have an influence over that particular joint. **Joint range of motion** is thus a functional measurement and not a measurement of the physical dimensions of the muscle or the length of the musculotendinous unit (Toppenburg & Bullock 1986). Measurements of the joint angle provide an ‘index’ of **musculotendinous unit length** or flexibility (Bullock-Saxton & Bullock 1994). For the purposes of this chapter, the term flexibility will be used to indicate an increase in joint range of motion.

Flexibility in the musculotendinous unit is directly related to tension in that musculotendinous unit (Starring et al. 1988; Taylor et al. 1990). Tension in the musculotendinous unit is made up of both a passive component and an active component (Taylor et al. 1990). The passive component lies in the connective tissue while the active component lies in the reflex activity of the muscle. The passive, connective tissue component has both viscous and elastic properties and therefore behaves in a viscoelastic fashion when subjected to a tensile force (Fung 1967; Ciullo & Zarins 1983; Herbert 1988; Taylor et al. 1990; Kwan et al. 1992).

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexibility</td>
<td>The range of motion in a joint or in a group of joints that is influenced by muscles, tendons and bones</td>
</tr>
<tr>
<td>Flexibility training</td>
<td>A training programme that is aimed at increasing the range of motion of specific joints or groups of joints (stretching programme)</td>
</tr>
<tr>
<td>Elasticity</td>
<td>The property in tissues that refers to the return of the tissue to its original length when the force is removed</td>
</tr>
<tr>
<td>Viscous</td>
<td>The property in tissues that refers to the elongation of a tissue that remains once the force applied to it is removed</td>
</tr>
<tr>
<td>Viscoelastic</td>
<td>A tissue that exhibits both viscous and elastic properties</td>
</tr>
<tr>
<td>Extensibility</td>
<td>The ability of a tissue to elongate</td>
</tr>
<tr>
<td>Stiffness</td>
<td>The ratio of the change in stress (force per unit area) and the change in length (strain)</td>
</tr>
<tr>
<td>Stress relaxation</td>
<td>The decline in tissue tension that results when a tissue is lengthened and then held at that constant length</td>
</tr>
<tr>
<td>Creep</td>
<td>The change in length of a tissue in response to a load that is applied and then maintained at that constant magnitude</td>
</tr>
<tr>
<td>Hysteresis</td>
<td>The change in length of a tissue in response to a load that is applied in a cyclic fashion</td>
</tr>
</tbody>
</table>
The elastic component of viscoelasticity implies that length changes in the musculotendinous unit are directly proportional to the loads imposed on the musculotendinous unit (Halbertsma & Goeken 1994). This is well described in Hooke’s model of a perfect spring (Burns & MacDonald 1975; Sapega et al. 1981; Taylor et al. 1990). Elasticity implies the ability of a structure to stretch when a force is applied to it, and then to return to the original length when the force is removed. Extensibility and stiffness comprise elasticity (Cavagna 1970; Halbertsma & Goeken 1994). Extensibility refers to the ability, in this case of the musculotendinous unit, to lengthen. Stiffness refers to the ratio of change in muscle length to the change in muscle tension (Halbertsma & Goeken 1994). Therefore, high stiffness implies a greater resistance to stretch. Thus, flexibility is comprised of two components—tension and stiffness (Fung 1967; Stromberg & Wiederheilm 1969; Herbert 1988).

In muscles, flexibility can also be defined as the point where tension first develops in the muscle when it is stretched passively. Stiffness is the ratio between tension and the amount of lengthening. Stiffness can therefore also be defined as the slope of the stress–strain curve (stress = tension/cross-sectional area; strain = change in length/initial length) (Burns & MacDonald 1975). The greater the stiffness, the steeper the slope of the stress–strain curve.

The viscous component of viscoelasticity is both rate change-dependent and time-dependent (Taylor et al. 1990). The rate of deformation is directly proportional to the applied force. This is illustrated by Newton’s hydraulic piston known as a dash pot (Burns & MacDonald 1975). The connective tissue in the musculotendinous unit responds as a viscoelastic material where the ground substances, collagen and elastin, which comprise the connective tissue, determine this response (Saagea et al. 1981; Starring et al. 1988). The relationship between length and tension, which changes with time as stretch is applied, implies viscoelasticity in the musculotendinous unit (Herbert 1988).

Stress relaxation occurs in viscoelastic substances when the tissue is stretched to a specified length and is then held at that constant length. The force or stress in the tissue, which will gradually decline over time, is known as stress relaxation (Taylor et al. 1990). In living soft tissue, this relationship is non-linear (Fung 1967; Herbert 1988; Taylor et al. 1990; Kwan et al. 1992). Generally, the elasticity of living tissue is characterized by a very small stress even when it is subjected to a large strain. When this stress begins to rise, however, it does so rapidly and exponentially (Fung 1967). Because the tension in the musculotendinous unit, when it is stretched, decreases over time, the tissue exhibits viscous behaviour, and because there is still a degree of tension present it also exhibits elastic behaviour (Taylor et al. 1990).

When the musculotendinous unit is placed under a constant force, the musculotendinous unit responds by lengthening in a viscous way. This implies an irreversible increase in musculotendinous unit length. The deformation in the musculotendinous unit occurs up to a point where the increase plateaus at a new length. This is known as ‘creep’ (LaBan 1962; Kottke et al. 1966; Fung 1967; Warren et al. 1976; Taylor et al. 1990). Prolonged, low load stretching of the musculotendinous unit results in more viscous deformation than short, high load stretching (LaBan 1962; Kottke et al. 1966; Fung 1967; Warren et al. 1976).

When a musculotendinous unit is stretched, it may not return to its original length or natural state because of its deviation from an ideal elastic substance (Starring et al. 1988). When the loads or stretches applied to the musculotendinous unit are cyclic, it becomes possible to predict the deformation to the musculotendinous unit. This is known as ‘hysteresis’ (Fung 1967). The effect of hysteresis is progressive, and further changes can be expected with repeated stretching (Starring et al. 1988).

Flexibility of the musculotendinous unit can, therefore, also be defined by the amount of extensibility and stiffness in that musculotendinous unit (Halbertsma & Goeken 1994). The degree to which the flexibility of the musculotendinous unit increases when a stress or stretch is applied...
to it, is determined by its viscoelasticity (Herbert 1988). The viscoelasticity of the musculotendinous unit allows for stress relaxation, creep and hysteresis to occur in that musculotendinous unit, when forces are applied at varying magnitudes, durations and repetitions (Kottke et al. 1966; Fung 1967; Warren et al. 1976; Starring et al. 1988; Taylor et al. 1990). Musculotendinous unit flexibility is also described as musculotendinous unit length and joint range of motion (ROM). The parameter, which is physically measurable, is joint ROM and this term will therefore be used to describe musculotendinous unit flexibility. An increase in joint ROM will indicate an increase in musculotendinous unit flexibility of the musculotendinous unit, which directly affects the joint ROM (Anderson & Burke 1991; Toppenburg & Bullock 1986).

General benefits of flexibility training

Athletes and trainers commonly perform flexibility training as part of regular training and in preparation for competition. It is also an integral part of most rehabilitation programmes following injury or surgery. The general benefits of flexibility training are usually cited as decreasing the risk of injury (primary and secondary prevention), decreasing muscle soreness and improving sports performance. Evidence that flexibility training does benefit the athlete will now be briefly reviewed.

Decreased risk of injury

Flexibility training is routinely recommended to increase ROM and thereby to reduce the risk of injury (Liemohn 1978; Glick 1980; Ekstrand & Gillquist 1982; Ciullo & Zarins 1983; Ekstrand et al. 1983; Zachazewski 1990; Beaulieu 1991). The rationale behind this recommendation is that musculotendinous units possessing greater flexibility are less likely to be overstretched, lessening the likelihood of injury (Bryant 1984). It has also been stated that increased flexibility decreases the risk of injury, because flexible joints can withstand a greater amount of ‘stress of torque’ before injury, compared to relatively stiff joints (Klafs & Arnheim 1981). It is important to consider independently whether increased flexibility reduces the risk of new injuries (primary prevention) or reinjury (secondary prevention). The scientific evidence for each of these two possibilities will now be briefly reviewed.

Does flexibility training decrease the risk of injuries?

A number of authors have proposed that decreased flexibility results in musculotendinous unit injuries (Liemohn 1978; Ekstrand & Gillquist 1982; Ekstrand et al. 1983; Worrell et al. 1991). It has been reported that hamstring flexibility was decreased in a group of injured athletes compared with a non-injured group participating in the same sport. Athletes with decreased ROM are more likely to suffer from musculotendinous unit tears than their flexible counterparts (Nicholas 1970; Worrell et al. 1991). It has also been suggested that muscle inflexibility may predispose athletes to certain injuries (Ekstrand & Gillquist 1982; Ekstrand et al. 1983). These investigations were conducted on soccer players, where the incidence of injuries was related to measures of flexibility. In a more recently published prospective study in military recruits, increased hamstring flexibility was related to a decrease in the overall incidence (% injuries in 13 weeks) (flexible group = 16.7%, control group = 29.1%) of injuries (Hartig & Henderson 1999).

However, there are studies that do not show that increased flexibility reduces the risk of injury. An early study on badminton players showed flexible players did not report less injuries than their inflexible team-mates (Henricson et al. 1983). More recently, in a well-conducted randomized prospective study, pre-exercise stretching did not reduce lower limb soft tissue, bony or all-injury risk in military recruits (Pope et al. 2000). However, the stretching protocol used in this study group (one static stretch of 20 s prior to training) may not have been sufficient to increase ROM.
Furthermore, a recently published comprehensive review of the literature has failed to produce strong evidence of a direct link to increased flexibility and a reduction in injuries (Shrier & Gossal 2000). Data from military populations seem to indicate that both increased flexibility as well as reduced flexibility is related to increased risk of injury (Jones & Knapik 1999). The relationship between injury risk and flexibility may well be a U-shaped curve with increased risk in both extremes (Fig. 11.1).

Despite these conflicting reports, stretching is still widely advocated and used by sportspersons and clinicians. In particular, the use of flexibility training to restore normal ROM following injury or surgery is still sound clinical practice. However, more studies are required to define the precise role of flexibility training in the prevention of injuries in sport.

**Decreased muscle soreness**

Muscle soreness, which follows vigorous or unaccustomed exercise, is known as delayed-onset muscle soreness (DOMS). Stretching has been advocated as a method to decrease DOMS (De Vries 1962; Liemohn 1978; Bryant 1984; Zachazewski 1990; Anderson & Burke 1991; Norris 1993). Despite the widespread belief that stretching decreases DOMS, there are few studies to verify this.

De Vries (1962) reported a chronic reduction of DOMS with repeated stretching. The total time of stretch was 5 min during the first 22 h post-exercise. Other reports of stretching to reduce DOMS are purely anecdotal (Noakes 1992). Some authors report that musculotendinous units that are flexible before being subjected to vigorous exercise, suffer less from DOMS than musculotendinous units that are inflexible before exercise (Noakes 1992).

In a study conducted by Buroker and Schwane (1989), subjects stretched for a total time of 45 min on the first day after exercise and for 25 min on the second and third days, respectively. McGlynn et al. (1979) stretched their subjects for a total time of 10 min on the first day and for 20 min on the second and third days, respectively. Neither of the studies found a decrease in DOMS with the addition of stretching. Buroker and Schwane (1989) mention that this may have been due to their stretch being held for too long.

None of the above results proves that stretching is effective in alleviating DOMS. It is clear that the conditions under which DOMS is decreased by stretching need to be investigated and defined.

**Improved performance**

Improved athletic performance may be another potential benefit of increased musculotendinous unit flexibility (Bosco & Komi 1979; Godges et al. 1989; Zachazewski 1990; Noakes 1992; Wilson et al. 1992; Norris 1993; Worrell et al. 1994). Indeed, in Borms (1984) it has been suggested that a limited ROM can reduce work efficiency. Two possible explanations are given for the positive association between increased flexibility and enhanced performance.

The first explanation is that musculotendinous flexibility is necessary in order for a limb to move through its full ROM. This is a given prerequisite for optimal performance (Ciullo & Zarins 1983; Borms 1984; Shellock & Prentice 1985). The second explanation is that flexibility training could reduce musculotendinous stiffness and enhance rebound performance (Wilson et al. 1992). Mechanical work is absorbed by the series elastic component of the musculotendinous
unit during eccentric contraction and is stored as potential energy. This energy is then available for use during the subsequent concentric contraction (Cavagna 1970; Bosco & Komi 1979).

Two studies have been conducted which investigate the effect of stretching on musculotendinous unit performance (Wilson et al. 1992; Worrell et al. 1994). Wilson et al. (1992) implemented a static stretching programme twice weekly for 8 weeks in the pectoral musculotendinous unit complex. At the end of this time, the bench press performance of the subjects had improved significantly. Worrell et al. (1994) determined the effect of hamstring stretching on hamstring musculotendinous unit performance. The subjects stretched 5 days a week for 3 weeks. Static stretching and proprioceptive neuromuscular facilitation (PNF) stretching were included as stretching methods. The results showed that both methods of stretching increased musculotendinous unit flexibility. Selective torque measurements were significantly improved for eccentric and concentric contractions of the hamstring musculotendinous unit in both groups.

It can be concluded that there is some scientific evidence that increased flexibility enhances musculotendinous unit performance. This is most probably because of the attainment of optimal ROM and the reduction of musculotendinous unit stiffness. Stored elastic strain energy is thereby utilized well during musculotendinous unit contractions.

Other benefits of stretching

Other reported benefits resulting from increased musculotendinous unit ROM are numerous. They include muscle relaxation (De Vries 1962), avoidance of skeletal dysfunction (Starring et al. 1988; Bullock-Saxton & Bullock 1994), prevention of scar tissue formation in the musculotendinous unit’s shortened position (Reilly 1992) and prevention of imbalance at the joints (Herbert 1988; Bullock-Saxton & Bullock 1994). Skeletal dysfunction can be defined as the syndrome that is a result of prolonged incorrect biomechanical usage of the joints that comprise the skeletal system (McKenzie 1985). Imbalance at the joints is the result of one musculotendinous unit at a joint being either significantly weaker or significantly less flexible than the other musculotendinous units at that joint (Kendall & McCreary 1983).

Summary

Numerous authors, trainers and athletes advocate the use of stretching to increase musculotendinous unit ROM. The benefits of an increase in musculotendinous unit ROM are commonly reported. These include decreased risk of musculotendinous unit injury, decreased DOMS, enhanced performance, avoidance of joint dysfunction, joint imbalance and scar tissue formation, and increased general relaxation and well being. There is, however, a paucity of scientific research data to substantiate these postulated positive effects of flexibility training.

There is some evidence that indicates that athletic performance may be enhanced by increased musculotendinous unit ROM. There is also an association between decreased injury prevalence and increased musculotendinous unit ROM. By all accounts, there are few, if any, negative side effects if stretching is implemented correctly. Therefore, the use of stretching to increase musculotendinous unit ROM is likely to continue to be supported by therapists, coaches and athletes.

Pathophysiology of soft tissue contracture following injury

The biology of soft tissue repair following injury or surgery has been discussed in detail in Chapters 3–5. However, in the context of flexibility training as part of normal rehabilitation following injury or surgery, a brief review of the mechanisms associated with soft tissue contracture is appropriate. A decrease in range of motion following injury and repair can be the result of a number of mechanisms (Table 11.2).

The healing process in soft tissues follows a continuous process consisting of four interwoven phases: immediate reaction to trauma (0–60 min), acute inflammatory phase (0–72 h), repair phase...
(2–42 days) and the remodelling phase (21 days to 1 year). During the late part of the inflammatory and early part of the repair phase, fibroblast accumulation is evident in the wound (Wahl & Renström 1990). The repair process follows and is characterized by recruitment of fibroblasts, proliferation of fibroblasts, collagen and matrix synthesis, and later, tissue remodelling.

Increased fibrosis within the scar

In normal tissue repair, matrix generation is self-limiting and minimal scar tissue formation occurs. However, in some wounds, scar tissue formation may become excessive. This process of increased fibrosis may result from: (i) increased numbers (proliferation and recruitment) of collagen-synthesizing cells; (ii) increased matrix synthesis by existing cells; (iii) deficient collagen degradation; and (iv) continued collagen synthesis (Wahl & Renström 1990). In addition to a larger volume of fibrous tissue, the fibrous tissue can also form adhesions and interfere with the normal function of adjacent tissue.

Wound contraction

Wound contraction refers to the inward movement of the wound edges by the action of cells known as myofibroblasts (Martinez-Hernandez & Amenta 1990). It is postulated that myofibroblasts are derived from mesenchymal stem cells or from perivascular cells. They appear at the wound site 3–4 days after the injury and become prominent 6–7 days after the injury. Their contraction can reduce wound size by up to 70%.

Excessive wound contraction can lead to contractions, which can result in decreased soft tissue flexibility. A detailed discussion on the control of wound contraction is beyond the scope of this chapter but has recently been reviewed (Nedelec et al. 2000).

Other factors

Following an injury, a decrease in musculotendinous unit flexibility may also be associated with other factors. These include: (i) increased nerve root tension (Gadjosik 1991); (ii) increased age (Norris 1993; Booth et al. 1994) and obesity (Norris 1993); (iii) decreased musculotendinous unit temperature (Rigby 1964; Sapega et al. 1981; Norris 1993); and (iv) immobilization or prolonged usage in a limited ROM (Herbert 1988; Starring et al. 1988).

Measurement of flexibility

The measurement of musculotendinous unit flexibility is of importance in both the laboratory and in the clinical examination room. A brief discussion on the basic method that is used to measure musculotendinous unit flexibility is therefore needed. Musculotendinous unit flexibility refers to its functional length (Toppenburg & Bullock 1986), and represents the degree to which it permits a movement over which it has an influence (Bullock-Saxton & Bullock 1994). In clinical practice, the ROM of a joint or a group of joints is usually measured and then compared with the normal values expected for a specific population. It is also useful to measure the ROM prior to starting flexibility training, in order to assess changes objectively. Practically, a clinical test of flexibility will be done by a clinician who will position a limb or a body segment, and then gradually move the segment to an end point.

For example, a number of anatomical positions have been used to assess hamstring musculotendinous flexibility. These include: (i) flexion at the hips in standing (Gadjosik et al. 1993); (ii) extension of the knee with the hip flexed to 90° while lying in the supine position (Gadjosik &

---

**Table 11.2** Mechanisms responsible for decreased range of motion in soft tissues following injury or surgery.

<table>
<thead>
<tr>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased fibrosis within the scar</td>
</tr>
<tr>
<td>Wound contraction</td>
</tr>
<tr>
<td>Immobilization</td>
</tr>
<tr>
<td>Nerve root tension</td>
</tr>
<tr>
<td>Decreased tissue temperature</td>
</tr>
<tr>
<td>Increased age</td>
</tr>
</tbody>
</table>

---

238 CLINICAL REHABILITATION INTERVENTIONS
Lusin 1983; Gadjosik et al. 1993); and (iii) the straight leg raise (SLR) method (Fahrni 1966; Breig & Troup 1979; Goeken 1991, 1993; Gadjosik et al. 1993; Cameron et al. 1994). The SLR method is commonly used because it is relatively easy to measure (Janda 1983; Kendall & McCreary 1983) and evidence indicates that it is a clinically acceptable form of measurement (Gadjosik 1991; Gadjosik et al. 1993). The SLR method has been proven as a reliable, albeit indirect, measure of hamstring musculotendinous unit flexibility (Gadjosik 1991; Gadjosik et al. 1993).

A number of methods have been used to determine the end-point for a clinical range of motion measurement. These include: (i) patient sensation of the end of range (as a tolerable, uncomfortable, but not painful sensation) (Goeken 1991); (ii) subjective tester determination of end-feel (Cyriax & Cyriax 1983; Maitland 1991); (iii) the start of joint movement as felt by the tester (Bullock-Saxton & Bullock 1994; Hanten & Chandler 1994); (iv) measurement of resistance to movement through objective force measures (Gadjosik 1991); (v) the objective use of cinematography to measure a limb angle (Bohannon 1982); and (vi) the use of the electromyogram (EMG) to objectively detect a sudden spike in neuromuscular activity (Moore & Hutton 1980; Bohannon 1984; Condon & Hutton 1987; Kirsch et al. 1993; Chequer et al. 1994). The repeatability of the various measures of end-point in testing hamstring flexibility was evaluated recently in a controlled laboratory-based study. The repeatability of the four common tests is depicted in Table 11.3.

The EMG detection of the end-point is the most repeatable measurement of the end-point of the flexibility test (intratester correlation coefficient = 0.99) (Hughes 1996), and is recommended for research studies. The most repeatable clinical method of assessing end-point is the end-feel by the tester (intratester correlation coefficient = 0.95). This method is recommended for testing end-range in a clinical setting.

### Physiology of stretching

Stretching implies that a force applied to the tissue elongates the tissue. The physiological effects that stretching has on the musculotendinous unit include neural effects (Beaulieu 1981; Reilly 1992; Norris 1993; Vujnovich & Dawson 1994), plastic (viscous) effects (Saepa et al. 1981; Starring et al. 1988) and elastic effects (Starring et al. 1988; Taylor et al. 1990). The physiology of each of these effects is discussed briefly.

#### Neural effects of stretching

There are three reflexes that primarily control the neural effects of stretching. These are the stretch reflex, inverse stretch reflex, and perception and control of pain by the Pacinian corpuscles.

When a stretch is applied to the musculotendinous unit, the first reflex evoked is the stretch reflex. The origin of this reflex is situated in the intrafusal and extrafusal fibres of the muscle spindle (Beaulieu 1981; Etnyre & Abraham 1986; Reilly 1992; Norris 1993; Vujnovich & Dawson 1994). The stretch reflex serves as a protective measure by causing a reflex contraction of the muscle to prevent overstretching of the musculotendinous unit. The magnitude and rate of contraction elicited by the stretch reflex in a musculotendinous unit are proportional to the magnitude and rate of the stretch that is applied to that musculotendinous unit (Beaulieu 1981; Matthews 1993). Both static and ballistic stretching evoke a response in the stretch reflex when the end-point of ROM is reached (Beaulieu 1981). One method of observing the stretch reflex is through the EMG device (Moore & Hutton 1980;
Condon & Hutton 1987). In this method, the stretch reflex causes a spiking of the neural activity, which is reflected by measuring EMG activity. The inverse stretch reflex applies both to prolonged contraction and stretching of the musculotendinous unit (Beaulieu 1981; Anderson & Burke 1991). This reflex is responsible for preventing prolonged increased strain of the musculotendinous unit, caused either by powerful active contraction or by overstretching of the musculotendinous unit (Beaulieu 1981; Etnyre & Abraham 1986; Anderson & Burke 1991). The origin of the inverse stretch reflex is located in the Golgi tendon organ (GTO), which is comprised of receptors situated in the musculotendinous junction of the musculotendinous unit. The GTO causes a dampening effect on the motor neuronal discharges, thereby causing relaxation of the musculotendinous unit by resetting its resting length (Beaulieu 1981; Reilly 1992; Norris 1993; Vujnovich & Dawson 1994). This usually occurs 6–20 s after commencement of the stretch (Norris 1993).

The third reflex of the musculotendinous unit originates in the Pacinian corpuscles that are located throughout the musculotendinous unit. The Pacinian corpuscles serve as pressure sensors, and assist with the regulation of pain tolerance in the musculotendinous unit (Beaulieu 1981; Reilly 1992; Norris 1993; Vujnovich & Dawson 1994). The three reflexes together are responsible for the neural effects that regulate musculotendinous unit flexibility, and are active during the stretching of the musculotendinous unit. The normal sequence of activation in a vigorous stretching movement would therefore be stretch reflex activation, causing reflex contraction of the musculotendinous unit, and Pacinian corpuscle activation, leading to a perception of pain. This would be followed in a prolonged stretch by GTO inhibitory effects and Pacinian corpuscle modification. The latter two reflexes will allow relaxation in musculotendinous unit tension and decreased pain perception (Beaulieu 1981; Reilly 1992; Vujnovich & Dawson 1994).

It has also been suggested that prolonged exposure to significantly large and unusual afferent inputs can cause long-term changes in levels of neuronal excitability and thereby, indirectly, on flexibility (Vujnovich & Dawson 1994). Therefore, musculotendinous units that are exposed to repetitive stretching demonstrate increased tolerance to the manoeuvre, resulting in an apparent increased ROM.

**Plastic and elastic effects of stretching**

Plastic deformation within the musculotendinous unit has largely been ascribed to changes which occur in the non-contractile elements within the musculotendinous unit (Cavagna 1970; Saepa et al. 1981; Vujnovich & Dawson 1994). The non-contractile elements (ligaments, tendons, capsules, aponeuroses and fascial sheaths) are predominantly comprised of connective tissue. Connective tissue is made up of collagen fibres which vary in spatial arrangements and are of varying densities and strengths (Saepa et al. 1981). The non-contractile element of the muscle tissue is also comprised of a connective tissue framework, which is responsible for its resistance to stretch (Saepa et al. 1981). Scar tissue and adhesions are also comprised of connective tissue (Saepa et al. 1981). Prolonged stretching results in plastic elongation of the connective tissue (Saepa et al. 1981; Starring et al. 1988). The mechanisms of both increased plastic and elastic deformation have already been defined and described above.

**Methods of stretching**

Several stretching techniques have been developed to increase joint ROM. These include PNF, ballistic stretching and static stretching. Combinations of these techniques have also been advocated. All these techniques rely on the neural, plastic and elastic effects of deformation as previously described for their effectiveness.

**Proprioceptive neuromuscular facilitation**

Proprioceptive neuromuscular facilitation is based on the theory that maximal contraction of
the musculotendinous unit results in maximum relaxation after active contraction of that musculotendinous unit (Knott & Voss 1956; Sady et al. 1982; Voss 1985; Osternig et al. 1987; Anderson & Burke 1991). The PNF stretches that are used in sports have been adapted from neurological physical therapy treatments. The two techniques that are commonly used are contract–relax (CR) and contract–relax agonist–contract (CRAC).

1 **Contract–relax.** In CR, a therapist stretches a limb to the point where limitation of ROM is felt. The therapist then holds this stretched position while the patient actively contracts the musculotendinous unit against the resistance of the therapist (Fig. 11.2).

The muscle is then relaxed and moved passively to the new lengthened position, until a stretch is felt in the new position (Hanten & Chandler 1994; Voss 1967, 1985; Norris 1993). A rotational component which entails a movement in the horizontal plane may be added to the resistance which is in the sagittal or longitudinal plane (Hanten & Chandler 1994; Voss 1967, 1985). Modifications to this method include contraction of the musculotendinous unit before stretching and contracting as described above (Beaulieu 1981; Taylor et al. 1990; Noakes 1992; Bandy & Irion 1994). The rationale behind the CR method is that increased stress through contraction of the musculotendinous unit will result in autogenic inhibition in the musculotendinous unit. The GTO activity will decrease tension and cause relaxation by resetting the musculotendinous unit length. This will facilitate the passive stretch procedure applied by the clinician (Norris 1993).

2 **Contract–relax agonist–contract.** In CRAC, the musculotendinous unit is stretched as described above in the CR procedure. In addition to this, the opposing musculotendinous unit (antagonist) of the musculotendinous unit being stretched (agonist) is contracted isometrically while the stretch is applied. The stretch procedure applied by the clinician is then aided by contraction of the opposing musculotendinous unit (antagonist). This results in reciprocal inhibition of the musculotendinous unit being stretched, in order to reduce its tension (Voss 1985; Norris 1993).

In one of the first studies on the role of PNF in flexibility training, PNF was compared with static stretching of the hamstring musculotendinous unit (Tanigawa 1972). In this study, subjects were divided into two treatment groups, and stretches were applied twice weekly for 4 consecutive weeks. The PNF stretch procedure involved taking the leg into SLR until the subject...
detected a stretch sensation. In this position, an isometric contraction was held for 7 s and then allowed to rest in the neutral position for 5 s. This was then repeated three more times. The static stretch procedure to the hamstring musculotendinous unit was applied to group 2 in the SLR position for 7 s and allowed to rest in the neutral position for 5 s. This static stretch was repeated three more times. The control group rested in the neutral position for 45 s. The results indicated that PNF was more effective than static stretching for increasing hamstring musculotendinous unit ROM. Static stretching was also more effective than the control group. Measurements were taken after each stretching session and these indicate that PNF resulted in greater increases in ROM at each of the individual stretching sessions and over the 4-week period.

In another study the effects of PNF, ballistic and static stretching were compared (Sady et al. 1982). Subjects were divided into four groups: PNF, ballistic stretch, static stretch and controls. Measurements of musculotendinous unit ROM were taken of the shoulder, hamstrings and trunk. These measurements were repeated at the end of the 6-week programme. Stretch procedures were performed three times weekly. The PNF procedure consisted of a stretch of the musculotendinous unit to end of range, an isometric contraction for 6 s, an unspecified relaxation period and a further unspecified stretch. This was repeated two more times. The ballistic stretch involved repeating the full ROM, rapidly, 20 times. The static stretch involved reaching the musculotendinous unit end of range, holding for 6 s, relaxing for an unspecified time and repeating the procedure twice more. The control group’s procedure was not specified. Results indicate that only PNF stretching increased musculotendinous unit ROM over the 6-week period, compared with the control group. This study was limited because there was no reported detail in terms of the control group’s procedure and individual measurements of musculotendinous unit ROM during the 6-week period.

Other authors also reported that PNF stretching techniques are effective in increasing musculotendinous unit ROM (Holt et al. 1970; Tanigawa 1972; Sady et al. 1982; Wallin et al. 1985; Etnyre & Abraham 1986; Osternig et al. 1987; Cobbing et al. 2001). However, if PNF is to be performed effectively, it requires expertise in all the stretching techniques, as well as an experienced clinician. These techniques, therefore, although effective, are seldom prescribed to or administered by individual athletes (Bandy & Irion 1994).

**Ballistic stretching**

During ballistic stretching, the limb is moved to the end of its ROM where the stretch sensation is felt, either passively by the clinician or actively by the subjects themselves (Fig. 11.3). Once this stretched position is achieved, repetitive bouncing or jerking movements are added (Schultz 1979; Sady et al. 1982; Hardy & Jones 1986; Anderson & Burke 1991). This jerking motion during ballistic stretching could theoretically exceed the limit of available range and thereby cause a strain injury or tear in the musculotendinous unit (De Vries 1962; Anderson & Burke 1991; Noakes 1992). In addition, this rapid jerking movement also elicits the stretch reflex, which then causes a reflex contraction against the direction of stretch in the musculotendinous unit. The resulting increase in tension can therefore increase the risk of injury (Beaulieu 1981; Anderson & Burke 1991; Norris 1993). Theoretically, this technique may not be as beneficial for increasing ROM as the musculotendinous unit would undergo a reflex contraction at each bounce or jerk due to the stretch reflex. Relaxation, or lengthening, of the musculotendinous unit can therefore not be achieved (Beaulieu 1981; Voss 1985; Noakes 1992; Matthews 1993).

In one study, the effects of a combination of ballistic and static stretching with static stretching alone on the ankle dorsiflexion musculotendinous unit were compared (Vujnovich & Dawson 1994). Subjects were divided into a static (stretch applied to the musculotendinous unit for 160 s) and a static–ballistic stretch group (static stretch applied to the musculotendinous unit for 160 s followed by 160 s of ballistic stretch). The
control group was rested for an unspecified time period. The results of this study showed that a static stretch followed by ballistic stretching increased musculotendinous unit ROM more than static stretching alone. Given that all the stretching techniques are reported to increase musculotendinous unit ROM (Tanigawa 1972; Sapega et al. 1981; Taylor et al. 1990; Vujnovich & Dawson 1994), it is not surprising that the addition of one technique to the other, thereby doubling the total stretching time, will result in greater increases in musculotendinous unit ROM than with one technique alone. In this study, the one group was stretched for 160 s in total, whereas the other group was stretched for a total of 320 s. The greater increases in ROM in the static–ballistic group could not therefore be explained by the stretch type only, but also by total duration of stretch.

The use of ballistic stretching would appear to increase the risk of injury due to the increases in musculotendinous unit tension. Furthermore, due to the stretch reflex mechanism, it, indeed, may not be effective in increasing musculotendinous unit ROM. Its use is therefore not widely advocated as the technique of choice to improve musculotendinous unit ROM.

**Static stretching**

The most common method of stretching that is used by athletes, coaches and therapists, is the static stretch (Zachazewski 1990; Anderson & Burke 1991; Beaulieu 1991; Norris 1993). Static stretching is characterized by the limb being moved slowly and gently to the end of its available ROM, in order to obtain a stretch sensation in the tissues (Fig. 11.4).

This controlled stretch is maintained in this position for an extended period (Beaulieu 1991; Norris 1993). The end of the available range is determined by maximum tolerance of the subject to the stretch sensation just short of pain (Norris 1993; Bandy & Irion 1994).

Static stretching is sometimes referred to as passive stretching. Where this occurs, the distinction referred to is the mode of application. If the stretch is applied and held by the subject, it is referred to as static stretching. If the static stretch procedure is applied by a therapist or another person to the subject, it is sometimes referred to as passive stretching (Beaulieu 1981; Noakes 1992). The danger of the passive stretch is that when executed incorrectly by a team-mate or partner, it has the potential to overstretch the...
musculotendinous unit. When administered by an experienced clinician, though, passive stretching can be very useful and as effective as the standard static stretch (Beaulieu 1981; Noakes 1992).

When the static stretch position is obtained slowly and gently, the stretch reflex response is dampened and occurs late in the movement. As the tension in the musculotendinous unit from the stretch and from the reflex contraction increases, the inverse stretch reflex initiated by the GTO is induced (Beaulieu 1981; Voss 1985). This causes autogenic inhibition in the musculotendinous unit and allows a greater ROM to be achieved in the musculotendinous unit (Voss 1985). The GTO, therefore, overrides the stretch reflex. In this way, both neural effects and plastic deformation are utilized to increase musculotendinous unit ROM.

In one study, variations of the static stretching procedures were examined. The effect of cyclic stretch with static stretch on hamstring musculotendinous unit ROM was studied (Starring et al. 1988). Subjects were divided into two groups and the hamstring musculotendinous unit length was measured using the knee extension test with the hip in 90° flexion. The cyclic stretch entailed the hamstring being stretched to the end of range and held for 10 s. The limb was then moved into a greater ROM for a further 10 s. This procedure was repeated for 15 min. Because the stretch was held statically for a period of time, this cyclic method of stretching is referred to as a repeated static stretch. The static stretch group had a constant force applied to the hamstrings for 15 min. Hamstring musculotendinous unit ROM was measured at the end of each treatment session. The procedures were repeated daily for 5 consecutive days. Another measurement was taken 5 days after the last stretching session.

Results showed that hamstring musculotendinous unit ROM increased more in the cyclic stretch group compared to the static stretch group. Increases in musculotendinous unit ROM were recorded on day 5 compared with day 1, for both groups. In both groups, there were still increases in musculotendinous unit ROM at 5 days after the last stretch session, compared with the measurements taken on day 1. These increases were, however, less on day 5 compared with those reported on day 1. The cyclic group maintained a significantly larger increase in ROM than the static group. These results show that static stretching which is held, and then repeated in the increased ROM, produces greater increases in musculotendinous unit ROM than static.
stretching sustained for long periods in one position.

Compared with other methods of stretching, static stretching produces the least amount of tension (Moore & Hutton 1980). Theoretically, it therefore has the lowest risk of injury, and is easy and safe to implement. The static stretching method is consequently the one most commonly advocated for use in a flexibility training programme (Bass 1965; Burry 1975; Glick 1980; Beaulieu 1981; Leach 1982; Zarins 1982; Voss 1985; Noakes 1992).

Scientific basis of a stretching session

A number of factors are important in the successful implementation of a stretching session or flexibility training programme, irrespective of the type of technique. These include: (i) the duration for which a stretch is held; (ii) the number of times the stretch is repeated in one session; (iii) the number of times the session is repeated per day or week; (iv) the effect of warm-up on the effect of stretching; and (v) the most effective anatomical position in which to stretch a particular musculotendinous unit. Despite the fact that there are guidelines in most sports medicine textbooks on the method of static and PNF stretching (Zachazewski 1990; Noakes 1992; Norris 1993), there are very few clinical research data available to support the currently advocated methods of static and PNF stretching. This section will focus on the recently acquired data from research studies that identified the scientific basis for optimum static and PNF stretching.

Optimum duration of a stretch at one stretching session

Static stretching

The duration of the force applied in static stretching can be defined as the time for which a stretch is applied at a single stretching session. Investigators have studied the effects of various durations of passive stretch in animal models. One study compared the effects of high load, short duration (5 min) stretch with low load (25% of high load), long duration (50 min) stretch (Warren et al. 1976). This study was conducted on rat tail tendon. The high load, short duration protocol resulted in almost twice the elongation compared to the low load, long duration group. However, in the high load, short duration group, elongation was retained for only half as long as in the low load, long duration group.

Lehmann et al. (1970) compared short-term stretch (time not reported) with long-term stretch (20 min) on rat tail tendon. This study found that long duration stretches were more effective in increasing tendon length. Both Lehmann et al. (1970) and Warren et al.’s (1976) studies did not investigate some of the commonly used stretching durations and therefore offer limited information on the optimum duration of static stretching. In addition, both these studies were conducted on animal tail tendon and therefore cannot be extrapolated to human musculotendinous units.

In one animal study, performed on rabbit musculotendinous units, stress relaxation was used to determine the effects of static stretch duration on musculotendinous unit ROM (Taylor et al. 1990). In this study, the musculotendinous unit was taken to a predetermined tension of 78.4 N (this magnitude of force was selected on the basis that it did not cause any irreversible damage but resulted in significant elastic deformation of the musculotendinous unit) and held at that length for 30 s. The stress relaxation was then measured by recording the change in tension in the musculotendinous unit over that time period. It was documented that significant relaxation in the musculotendinous unit occurred after 12–18 s, although relaxation continued over the entire 30 s period. Because the measurements were only taken over 30 s, it is impossible to state whether the same overall stress relaxation would have occurred if held for a period of 48–72 s (i.e. 4 × 12–18 s).

The optimum duration of a static stretch that is to be used in the normal population, and in athletes in particular, has also been investigated in human studies. The optimum duration of static stretching has been reported on numerous
occasions and varies from 6 s to 20 min (De Vries 1962; Tanigawa 1972; Medeiros et al. 1977; Moore & Hutton 1980; Beaulieu 1981; Klafs & Arnheim 1981; Sady et al. 1982; Bohannon 1984; Henricson et al. 1984; Madding et al. 1987; Zachazewski 1990; Bandy & Irion 1994).

Until recently, only three studies have been reported in humans where different durations of static stretch were compared with another. Bohannon (1984) measured increases in musculotendinous unit length at 15 s and then at 30 s intervals thereafter, for 8 min. Madding et al. (1987) measured only at 15, 45 and 120 s. Bandy and Irion (1994) measured at 12, 30 and 60 s over a 6-week period.

A definitive study described by Hughes et al. (2001) examined the effects of different durations of static stretching on hamstring musculotendinous unit ROM. In this study, 10 normal subjects underwent static stretching durations of 10, 20, 30, 60, 90, 120 and 240 s. ROM was measured using EMG activity as the end-point. Three stretches were applied with a 1 min rest in between. The results of this study are depicted in Fig. 11.5.

The results of this study clearly show that a static stretch duration of 30–60 s is optimal to increase ROM in static stretching.

The precise duration of a PNF stretch on increases in hamstring ROM has only recently been reported. In this study, human subjects (experimental and a control group) were subjected to durations of PNF stretch of 10, 20, 30 and 60 s. The increases in hamstring ROM following different durations of PNF stretches are depicted in Fig. 11.6.

The results of this study show that a 10 s PNF stretch is as effective in increasing ROM as longer durations, i.e. up to 60 s.

Number of repetitions (frequency) of stretches at one stretching session

The optimum number of repetitions that should be performed at a single static stretching session has not been investigated until recently. In the past, between one and 20 repetitions have been advocated to increase ROM (Beaulieu 1991; Norris 1993). However, this advice has not been based on scientific data from human studies.

In one animal study on rabbit musculotendinous units, this parameter was investigated (Taylor et al. 1990). In this study, the stress–relaxation
procedure was used. This involved taking the musculotendinous unit to a predetermined 78.4 N point of tension. This point was determined as the point where no irreparable damage was done to the musculotendinous unit, but where lengthening occurred. The length of the musculotendinous unit where this point of tension occurred was maintained for 30 s. During the 30 s period, the tension in the musculotendinous unit and its length was measured. The tension decreased during the 30 s period. After 30 s, the musculotendinous unit was returned to the initial resting tension of the musculotendinous unit of 1.96 N. This cycle of stretching was repeated 10 consecutive times for each musculotendinous unit. Results showed significant increases in musculotendinous unit length during the first four cycles, with cycles 1 and 2 showing the largest increments in increased musculotendinous unit length. Cycles 5–10 also showed increases in musculotendinous unit length, but these were not as significant as those occurring during cycles 1–4.

This study indicates that the optimal number of repetitions in a single stretching session is four. However, because all the stretches were only held for 30 s, it was not possible to determine whether the same increases in ROM would have been attained had the stretch been held for a total of 48–72 s (i.e. 4 × 12–18 s) in one session. Although this was a very thorough and useful study, it required validation in the human model.

It was therefore necessary to investigate the number of repetitions in a static stretch session that are required to gain optimum increases in ROM in the human musculotendinous unit. Recently this study was undertaken, using normal subjects that were subjected to stretches of 30 s duration (Hughes 1996). The frequency of stretches varied from one to 10 stretches. The results of this study are depicted in Fig. 11.7.

The results of this study indicate that three stretches are optimum in increasing ROM and that additional stretches do not add any further advantage to increasing the ROM.

Fig. 11.7 The effect of different static stretch repetitions on changes in hamstring ROM. All stretch repetitions were significantly different from controls. *, significant difference from change in ROM from after repetitions. (From Hughes et al. 2001.)

PNF STRETCHING

The optimum frequency of PNF stretches to increase hamstring ROM has only very recently been reported. Healthy subjects underwent 10 s repeated PNF stretches of different frequencies. Hamstring ROM was assessed in a stretch and control group using EMG spikes as the end-point to determine ROM. The results of this study are depicted in Fig. 11.8.

The results of this study show that three PNF stretches result in the largest increase in ROM, but that this was not significantly more than performing a single stretch.
Retention of increased ROM after one stretching session

**STATIC STRETCHING**

Until recently, it had not been well documented how long the retention in increased ROM following a static stretch session would last. A number of authors have suggested time periods for the retention of ROM, but none of these have been tested in a well-conducted clinical trial. Suggested time periods for retention of ROM ranged from 3 h (Beaulieu 1981) to days (Atha & Wheatley 1976). Data on the retention of ROM after a static stretching session are vitally important in order to implement a scientifically based stretching programme.

In a study on rat tail tendon by Warren et al. (1976), it was demonstrated that the application of low load forces (25% of high load) for 50 min resulted in the retention of more than twice the elongation compared with the high load, short duration applied for 5 min. The retention of elongation was measured 10 min poststretch. This study was conducted on animal tendon and is therefore not representative of the complete musculotendinous unit and may also not be applicable to human subjects.

In a recent study (Hughes 1996), the retention in ROM achieved by a static stretch session (3 × 30-s stretches) was evaluated. Normal subjects underwent a static stretch session, and ROM was then measured over a 24 h period to determine residual increases in ROM at periodic intervals. The results of this study are depicted in Fig. 11.9.

The results of this study indicate that following a stretch session that is conducted optimally, based on recent scientific evidence, ROM is only retained for 4–6 h. The practical implication is that the frequency with which stretch sessions should be conducted during a day to achieve maximal benefit, is approximately every 6 h (during awake hours).

**PNF STRETCHING**

The retention in ROM following a PNF stretch session has not been well investigated. In one study by Toft et al. (1989), it was documented that there was a residual effect of increased musculotendinous unit ROM for at least 90 min after a PNF stretching bout. In this study, measurements of ankle dorsiflexion were taken. PNF stretching was obtained by maximal contraction of the plantar flexors for 8 s in the stretched position, followed by 2 s of relaxation and finally 8 s of slow static stretching. This procedure was repeated five times, and 90 min later ankle dorsiflexion was remeasured. It was found that the increases in musculotendinous unit ROM were maintained over this 90 min period. Here the static stretching procedure was administered for 40 s only in total (15 × 8 s) and was preceded by an isometric contraction before each stretch. The duration of retention of ROM following PNF stretching requires further investigation.

**Cumulative effect of repeated static stretching sessions**

**STATIC STRETCHING**

Until recently, there had been no studies that documented the cumulative effect, if any, of
repeated bouts of static stretching on joint ROM. In one study by Bandy and Irion (1994), subjects were placed on a stretching programme for 3 weeks. Measurements of hamstring musculotendinous unit ROM were made before and after the study. Subjects in the three experimental groups performed one 15 s, one 30 s or one 60 s static stretch. This static stretch was performed daily, 5 days a week, for 3 weeks. At the end of this period, hamstring musculotendinous unit had increased significantly in the groups stretching for 30 and 60 s, but not in the group stretching for 15 s. However, because measurements were only taken prior to the onset of the study and after 3 weeks, it was not possible to determine when the changes in ROM occurred.

More recently, the cumulative effects of static stretching were investigated. In this study healthy subjects underwent daily stretching every six hours (3 × 30 s static stretches). ROM measurements of the hamstring musculotendinous unit were conducted over a 22-day period. The results of this study are depicted in Fig. 11.10. The results of this study indicate that daily static stretching (every 6 h) can result in retention of ROM after as short a period as 7 days. There is also a progressive increase in ROM (approximately 1% per degree, in general). From Fig. 11.10 it can be seen that the increase in ROM achieved using this stretching programme is approximately 15–20%, and that a plateau effect had not yet been achieved after 3 weeks.

PNF STRETCHING

Retention in ROM following PNF stretching has not been well investigated either. In one study, a PNF stretching programme was implemented where subjects stretched twice daily for 3 weeks (Toft et al. 1989). Ankle dorsiflexion ROM was measured before implementing the stretch procedure and after 3 weeks of stretching. The stretch procedure consisted of a maximal isometric contraction of the plantar flexors for 8 s in the stretched position, followed by a 2 s rest period and finally an 8 s slow static stretch of the plantar flexors. This was repeated five times. The whole procedure was carried out twice daily for 3 weeks. No measurements of ankle ROM were taken during the 3-week period. Ankle ROM was remeasured at the end of the 3-week period. At the final measurement, no stretching procedures were carried out for at least 20 h prior to measurement. There were still large residual increases in ankle musculotendinous unit ROM.

Role of warm-up prior to static stretching on musculotendinous unit ROM

The inclusion of warm-up prior to stretching is widely advocated (Dominguez 1982; Wathen 1987; Zachazewski 1990). However, its role in altering the effectiveness of the stretching procedure is not clear. It has been suggested that increases in flexibility occur as a result of
increases in tissue temperatures (Cureton 1941; De Vries 1962; Lehmann et al. 1970; Nicholas 1970; Warren et al. 1976). Other authors suggest that increased tissue temperatures do not result in increases in musculotendinous unit flexibility (Halkovich et al. 1981; Wiktorsson-Moller et al. 1983; Williford et al. 1986; Cornelius et al. 1988).

It is also important to determine whether the warm-up is passive or active. Passive warm-up occurs when the muscles are warmed by means of external heating such as direct heat application (heat packs) or in the form of therapeutic ultrasound. Active warm-up refers to a warm-up of the muscles by active contraction. This would be in the form of a jog or any other exercise where the muscles to be stretched are repeatedly contracted prior to stretching.

**Passive Warm-up**

It has been postulated that in passive warm-up, the increases in tissue temperatures increase the extensibility to the musculotendinous unit. This increased extensibility is due to the effects of heat on the intermolecular bonding of collagenous tissue. It is probable that heating of the musculotendinous unit results in the collagen being partially destabilized and consequently more pliable and extensible (Rigby 1964; Lehmann et al. 1970; Warren et al. 1976; Saapega et al. 1981; Ciullo & Zarins 1983).

In an animal study, Warren et al. (1976), showed that rat tail tendon extensibility was increased with the addition of heat prior to the static stretch procedure. The tendons were either exposed to temperatures of 25, 39 or 45°C. A load of 100 g was applied to all the tendons and increases in length were measured. The groups exposed to higher temperatures experienced significantly greater increases in extensibility compared with the 25°C group at the load of 100 g. There was also less resultant tissue damage in these two groups than in the 25°C group.

Studies on the effects of passive warm-up on static stretching have also been performed on human subjects. In 1984, Henricson et al. conducted a trial on the effects of heat and stretching on musculotendinous unit ROM. Measurements of hip flexion, abduction and external rotation were taken. Subjects were divided into three groups. The first group had an electric heat pad applied for 20 min at 43°C. A second group had the electric heat pad applied as above and this was then followed by a modified PNF technique consisting of 7 s isometric contraction of the musculotendinous unit, followed by 7 s relaxation, followed by 7 s static stretching. The third group performed only the PNF technique of stretching. Results show that no increases in ROM occurred in the first (heating only) group. The other two groups showed significant increases in ROM, with the combined heat and PNF group showing the greatest increases in ROM. This study indicates that passive heating alone is not effective in increasing musculotendinous unit ROM. However, when combined with stretching, passive heating is more effective in increasing musculotendinous unit ROM than stretching alone. The PNF stretching technique is notable in that it includes a muscle contraction as well as a static stretch. Therefore, PNF techniques employ an element of active warm-up as well.

In another human study, the effect of passive heating on the stretching procedure was studied (Wessling et al. 1987). In this study, the triceps surae musculotendinous unit either had static stretching applied to it for 1 min, or 1 min of static stretching applied after 7 min of continuous ultrasound. Ultrasound causes a deep heating effect in tissues (Gerston 1955; Low & Reed 1990). This study showed that the addition of deep heating to static stretching of the musculotendinous unit resulted in greater increases in ROM than static stretching alone. The difference between this study and the one conducted by Henricson et al. (1984) may lie in the deep heating effect of ultrasound compared with the more superficial heating of the heat pad.

**Active Warm-up**

Although knowledge of the effects of the addition of passive heat to the effectiveness of stretching in increasing musculotendinous unit ROM is
valuable to the clinician, active warm-up is more practical and of more value to the sportsperson. As muscles contract repeatedly, heat is produced and the intramuscular temperature increases (Zachazewski 1990). This process is known as active warm-up or physiological preconditioning (Safran et al. 1988; Zachazewski 1990). There have been a number of studies conducted to determine the effect of active warm-up on musculotendinous unit ROM.

In an animal study conducted by Safran et al. (1988), the effect of active warm-up on the ROM of the musculotendinous unit was investigated in rabbits. In one group, musculotendinous unit length was measured and then this musculotendinous unit was electrically stimulated to its maximum wave summated tension for 15 s. This was classified as the active warm-up period. The electrode was then removed and the musculotendinous unit was statically stretched to failure. In the second group the musculotendinous unit length was measured and then statically stretched to failure without active warm-up. An average of 7% more force was needed to reach failure in the active warm-up group. The actively warmed-up musculotendinous units could also be stretched to a significantly greater length before reaching failure. Intramuscular temperatures were recorded to increase by 1°C in the actively warmed-up musculotendinous units. The results of this study are valuable because they indicate that actively warmed-up musculotendinous units obtain significant increases in musculotendinous unit ROM and can absorb significantly higher forces before failure. This would indicate that actively warmed-up musculotendinous units would significantly increase their musculotendinous unit ROM when statically stretched and also have a lower risk of injury.

In another study, the effects of active warm-up, massage and stretching on ROM and muscle strength were conducted (Wiktorsson-Moller et al. 1983). In this study, hip flexion, hip extension, hip abduction, knee flexion and ankle dorsiflexion ROM, and hamstring and quadriceps strength were measured. There were four groups that were compared. The first group received an active warm-up consisting of 15 min of stationary cycling. The second group received 15 min of warm-up as above and 12 min of massage. The third group received 12 min of massage only. The fourth group received warm-up as above and PNF stretching. The PNF stretch consisted of a maximal isometric contraction for 4–6 s, followed by relaxation for 2 s and finally by a static stretch for 8 s. Results showed that other than for ankle dorsiflexion, warm-up alone, massage alone and warm-up plus massage have no effect on musculotendinous unit ROM. Warm-up combined with stretching showed significant increases in ROM for all joints. None of the groups showed any changes in muscle strength with the intervention. No measurements were done on the effects of stretching alone, and therefore the effect of the addition of warm-up to stretching could not be determined.

In 1986, Williford et al. conducted a study on the effects of active warm-up on musculotendinous unit flexibility over a period of 9 weeks. The flexibility of the shoulder, hamstrings, trunk and ankle were measured. The subjects were divided into two groups. The first group performed only static stretching that entailed a static stretch, which was held for 30 s and then repeated once more, totalling 60 s of static stretching. The second group actively warmed-up by jogging lightly but progressively for 5 min, prior to performing the same static stretching as in group 1. The subjects performed their respective stretching routines twice per week for 9 weeks. Results indicated that the active warm-up group gained significantly greater increases (approximately 14%) in ROM than the stretching only group. Gains (approximately 11%) in ROM were also documented in the stretching-only group.

In another study, the effects of the placement of a stretching session in a work-out in increasing ROM was determined (Cornelius et al. 1988). Subjects were divided into three groups. The first group performed static stretching for 10 s before the exercise session, which consisted of 20–30 min of jogging. The second group performed static stretching after the exercise session described. The third group performed the
stretching both before and after the exercise session. The stretching and exercise sessions were performed 3–5 times per week, for 6 weeks. Results showed increases in musculotendinous unit ROM in all the groups, but no significant differences between the groups. This study is limited in its value because the stretching procedure was not well defined, and the subjects in the groups were not well matched. There was a wide variation in the amount of stretching and exercise that was performed by the subjects.

From the above studies, it is difficult to determine the effects of active warm-up on musculotendinous unit ROM. Passive warm-up appears to have little effect on musculotendinous unit ROM (Henricson et al. 1984). When passive warm-up is combined with stretching it appears to be more effective than stretching alone for increasing musculotendinous unit ROM (Henricson et al. 1984; Wessling et al. 1987). Active warm-up before static stretching in human subjects appears more beneficial in increasing musculotendinous unit ROM than static stretching alone, over a period of 9 weeks (Williford et al. 1986). The most valuable study on the effect of active warm-up was conducted on rabbit musculotendinous units (Safran et al. 1988). However, the results of this study cannot be directly extrapolated to humans. In this study, active warm-up increased musculotendinous unit length to failure, and the force necessary to reach failure.

In a recent study, the effects of active warm-up alone, static stretching alone and active warm-up combined with static stretching on musculotendinous unit ROM were determined. Normal subjects underwent either active warm-up only (running until 70% maximum heart rate was achieved), static stretching (3 x 30 s) only, or combined active warm-up and then static stretching. Hamstring ROM measurements were conducted in all groups after the interventions, and in a control group. The results of this study are depicted in Fig. 11.11. The results of this study indicate that active warm-up alone is not effective in increasing ROM, but that it may have an additive effect if it is performed prior to static stretching.

There have been no studies that compare active versus passive warm-up, with and without stretching.

**Effect of stretch position during the stretching session on musculotendinous unit ROM**

It is a widely held belief that stretch position alters the efficacy of the stretching procedure (Voss 1985). For example, it is generally believed that increased ROM in a musculotendinous unit cannot be achieved when the musculotendinous unit is contracting. In a recent study, the effect of limb position on hamstring ROM has been evaluated. Normal subjects underwent three types of stretching procedures: (i) standing and stretching a weightbearing (contracting) hamstring muscle; (ii) standing but stretching a non-weightbearing hamstring muscle; and (iii) lying supine and stretching a non-weightbearing...
Hamstring muscle. The ROM immediately after the stretches (3 × 30 s static stretches) was measured. The results are depicted in Fig. 11.12.

The results of this study show that limb position is important when performing static stretch procedures. It is advised that static stretching should take place in the relaxed, non-weightbearing position.

Conclusions and practical guidelines for stretching

- Musculoskeletal injuries are common in athletes and require rehabilitation following injury or surgery.
- Injury or surgery can result in decreased joint ROM mainly due to fibrosis and wound contraction.
- Flexibility training is an important component of rehabilitation in order to minimize the decrease in joint ROM.
- Flexibility training should begin in the first week following injury or surgery to prevent decrease in ROM due to wound contraction and fibrosis.
- The main postulated benefits of flexibility training for the non-injured athlete are injury prevention (weak scientific evidence), increased athletic performance (good scientific evidence) and the prevention of muscle soreness (some scientific evidence).
- A variety of stretching techniques can be used in improving range of motion, including PNF, ballistic stretching and static stretching.
- Static stretching is most commonly used, is safe and effective.
- PNF is more effective than static stretching but it requires some training to master the technique and requires an assistant in some instances.
- Ballistic stretching may be effective but it has been postulated to increase the risk of injury while performing the stretch.
- An optimum static stretch session should consist of three stretches lasting 30 s each.
- Two to three static stretch sessions should be performed daily for maximum benefit.
- Static stretch sessions performed two or three times per day will result in a permanent increase in joint range of motion after 7 days.
- A cumulated benefit to daily static stretches (every 6 h) for up to 21 days has been demonstrated.
- A warm-up alone will not increase ROM but has an additive effect if it precedes a static stretch session.
- A static stretch should be performed in the non-weightbearing position with the muscle relaxed for maximum gain in ROM.

References


Lower extremity muscular flexibility in long distance 
rinters. *Journal of Orthopaedic and Sports Physical 

Heat and stretch procedures: an evaluation using rat 
tail tendon. *Archives of Physical Medicine and Rehab-

Wathen, D. (1987) Flexibility: its place in warmup activ-
ities. *National Strength and Conditioning Association 

Weber, S. & Kraus, H. (1949) Passive and active stretch-

Effects of static stretch versus static stretch and ultras-
ound combined on triceps surae muscle extensibility 

Wiktorsson-Moller, M., Oberg, B., Ekstrand, J. et al. 
(1983) Effects of warming up, massage, and stretch-
ing on range of motion and muscle strength in the 
lower extremity. *American Journal of Sports Medicine* 
**11** (4), 249–252.

Williford, H.N., East, J.B., Smith, F.H. et al. (1986) 
Evaluation of warm-up for improvement in flexibility. 

shorten cycle performance enhancement through flexi-
bility training. *American Journal of Sports Medi-
cine* **24** (1), 116–123.

(1991) Comparison of isokinetic strength and flexi-
bility measures between hamstring injured and non-
injured athletes. *Journal of Orthopaedic and Sports 

Effect of hamstring stretching on hamstring muscle 

In: *Sports Physical Therapy*. Appleton & Lange, East 
Norwalk, CT: 229–230.

bimechanical aspects. *International Journal of Sports 
Chapter 12

Strength and Endurance

GUNNAR GRIMBY AND ROLAND THOMEÉ

Introduction

Injuries to the musculoskeletal system could result in skeletal muscle hypotrophy and weakness, loss of aerobic capacity and fatigability. Often there is a combination of these negative effects of injuries and immobilization with usually one or the other dominating, depending on the type of injury. This chapter will discuss the time sequence and extent of these changes in the perspective of rehabilitation measures. Rehabilitation techniques and exercise prescription will especially be described for resistance training and cardiovascular (aerobic) conditioning training during the second phase (after the acute phase with immobilization).

Reduction in skeletal muscle strength

Reduction in skeletal muscle strength and increased fatigue after injury and/or immobilization with disuse depend on several factors (Kannus et al. 1992b,c; Behm & St Pierre 1998). The specific effects of immobilization have been described in detail in Chapter 6 and the physiological and functional implications of injury in Chapter 7. In the clinical management nowadays, immobilization is avoided, if possible. The reduction in muscle strength can in principle depend on impaired neuromuscular activation and/or reduced muscle volume, as it can be assumed that the number of muscle fibers are protected.

Impaired muscular activation

An impaired muscular activation can depend on factors such as: (i) inhibition due to pain or other afferent nerve influence; or (ii) reduced muscle activity at immobilization and/or injury.

Pain, joint effusion and knee angle are factors influencing the degree of inhibition. Pain from a joint may result in involuntary inhibition of activity in muscles acting across that joint, and a considerable lack of ability for muscle activation has been seen after knee pathologies and surgery (Hurley et al. 1994). However, the presence of pain is not necessary for inhibition to occur (Stokes & Young 1984; Shakespeare et al. 1985). There is not a direct relationship between the degree of pain and the amount of muscle inhibition. Inhibition may be as severe in a later postoperative period as at the first postoperative day, even in a virtually pain-free state. The activity in other receptors from the joint and adjacent structures may have similar effects as seen in animal experiments and human studies. It has been demonstrated that a very small amount of joint effusion will result in reflex inhibition of the musculature. The amount of quadriceps inhibition after knee injury may partly be dependent on the knee angle, with less inhibition in a flexed than in a fully extended position (Shakespeare et al. 1985). There might be selective inhibitions of motor units, which could explain part of the dominant hypotrophy of a specific fibre type (see below), seen after joint damage. Selective
inhibition of part of the quadriceps muscle has also been noted among patients with patellofemoral pain syndrome (Thomée et al. 1996).

There is a vicious cycle causing muscle weakness. Figure 12.1 illustrates the relationship between joint damage, immobilization, reflex inhibition and muscle wasting. The reflex inhibition can therefore result in an immediate reduction of voluntary muscle activation causing muscle weakness, early after an injury or surgery. In a longer perspective it will also cause muscle wasting resulting in further muscle weakness.

Reduced muscle volume

The other factor causing decreased muscular strength, besides inhibition, is the reduction in muscle volume that follows an injury. In rehabilitation, aspects to consider are the importance of the different patterns of muscle hypotrophy, the difficulty in assessing the degree of muscle wasting, the difference in strength between extremities, and the strength relation between the agonist and the antagonist muscles (agonist–antagonist ratios).

Patterns of muscle hypotrophy

Different patterns of muscle hypotrophy may be seen depending on the type of immobilization and disuse (Appell 1990; Kannus et al. 1992b,c). Patients with longstanding pain may show hypotrophy of slow-twitch type I fibres, as also seen after immobilization with a short muscle length. After general disuse, a reduction, especially of fast-twitch type II fibres, predominates. A reduction of the muscle size of the quadriceps equivalent to up to 30–40% may be seen after knee injury. The wasting of quadriceps usually dominates over the loss of hamstrings muscle volume.

Assessing the degree of muscle wasting

The degree of muscle wasting can not be accurately diagnosed by measurements of circumference using a standard tapemeasure device and can in fact be grossly underestimated. The reason for this is that the fat and bones structures are practically constant. With a reduction of the circumference of the muscle tissue, there will be an increase in the width of the subcutaneous fat layer masking part of the effect of muscle wasting on the total circumference. Thus, it is difficult and often misleading to use changes in circumferences as indicators of gains in muscle tissue during rehabilitation after injury or immobilization. Other more exact means have to be used including computed tomography (CT), magnetic resonance imaging (MRI) or ultrasound. In clinical practice one has often to rely on the indirect measurement of muscle strength and the interpretation of changes in muscle strength with respect to muscle volume.

Difference in strength between extremities

An important aspect when evaluating reduction in strength at follow-up, is that there may be a lack of perception of a moderate difference in strength between the two extremities. Thus, a 20% reduction, or even more, of quadriceps muscle strength has been recorded in the previously injured side compared with the non-injured side in persons who have returned to normal activity and considered themselves fully recovered (Grimby et al. 1980; Arvidsson et al. 1981). If such differences are seen in athletes, objective evaluation and further rehabilitation are needed.
**Agonist–antagonist ratios**

Assessment of the ratio of the agonist–antagonist strength, as the quadriceps–hamstrings ratio, may also result in recommendations for further training, but a comparison between injured and non-injured sides is of more importance in clinical practice. The knowledge of the prevalence of muscular imbalance and its functional significance is, however, rather poor.

**Reduction in muscle endurance**

Parallel with the reduction in muscle strength there will usually be a reduction in muscle endurance. A given level of force production will now be closer to the maximal force output of the muscle, and therefore will need a relatively higher proportion of the muscle strength. There are a number of factors which may reduce muscle endurance following disuse (Kannus et al. 1992b, c; Behm & St Pierre 1998), including a reduction in mitochondrial enzymes necessary for the oxidative process in the muscle and a reduction in capillarization with less ability to transport oxygen and energy supply to the muscle cell. On the other hand, a reduction in muscle fibre cross-sectional area with disuse will reduce the diffusion distances. With training these factors will be reversed. An increase in mitochondrial enzymes will be seen quite quickly during adequate training emphasizing the endurance adaptation.

**Background for resistance training**

There are different aspects of resistance training depending upon whether an improved neuromuscular activation or increase in muscle volume are given priority.

**Neuromuscular activation**

Important points to consider are time, assessment of the degree of neuromuscular activation, contralateral effects of resistance training, and speed of force development. The specific nature of the neural adaptation is relatively unknown and different aspects may be of importance, depending on the type of training, the muscles involved, and the pretraining status.

The mechanisms of adaptation may include increased activity of the prime movers for a specific movement and changes in the activation of synergists and antagonists (Sale 1988). There may be an increased activation of the number of motor units at maximal effort and/or an increased firing rate. The relative importance of these factors may vary between muscles and type of muscle activation. Motor units may be better synchronized. There may be facilitation of reflexes as well as reduction of inhibitory influence. Parallel to the activation of the prime mover, activation of the antagonist may also occur, having a functional positive or negative effect. A reduction in the activation of the antagonist during training may lead to greater activation of the prime mover, reducing any reciprocal inhibition, and also—through less force produced by the antagonists—an increased net output of force in the direction of the movement.

**Time aspects**

Early after an injury and/or immobilization there is most likely to be reduced neuromuscular activation (Sale 1988; Häkkinen 1994). Strength training effects will then mainly depend on neural factors and may be limited to the type of training exercise, both with respect to speed during dynamic exercise and joint angle during static exercises (Kraemer et al. 1996). The length of this initial period may vary depending on the type and degree of injury and disuse. For training from an initial moderately untrained state, it has been estimated to be in the order of 6–10 weeks. The initial training effect may be quite large, providing that there are no major inhibitory factors, as discussed earlier.

**Assessment of neuromuscular activation**

One way in the laboratory setting to identify lack of neuromuscular activation is the use of single superimposed twitch electrical stimulation (Behm
After identifying the maximal intensity of the stimulus for a single twitch by successively increasing the signal during resting conditions to get a plateau, one or several single-twitch stimuli are given during muscle actions with maximal voluntary effort. Electrical stimulation has mainly been used during static muscle activity, whereas our knowledge is more limited concerning optimal neuromuscular activation during dynamic muscle actions. During these conditions, electrical stimulation with tetanic trains has been used; however, this requires considerably cooperation from the subject due to the discomfort involved. It has been used to demonstrate the lack of maximal activation at eccentric muscle actions (Westing et al. 1990).

**Contralateral effects of resistance training**

It has been reported that by training one leg, the other leg also may improve and show strength gains. Improvements can be in the order of 35–60% of that attained in the trained limb (Kannus et al. 1992a). Such a possible effect of contralateral training should be recognized and possibly be used during rehabilitation after injury and immobilization of one extremity.

**Speed of force development**

Another aspect of muscle force development is a shortening of the time for tension development after activation (electromechanical delay). Speed of force development may decrease after immobilization or surgery (Komi 1984) and can increase with specific training. The force–time relationship may change using power training with high-velocity contractions or explosive training models, so that there will be a relatively faster force rise early in the muscle activation (Häkkinen 1994).

**Muscle volume**

As mentioned earlier, in the initial stages of weight training for sports or rehabilitation purposes, increases in strength occur mainly because of neural adaptation. In contrast, in the later stages of resistance training increases in muscle strength occur mainly by increases in muscle volume (Fig. 12.2) (Sale 1988). There is an increase in the size (cross-section) of the individual muscle fibres leading to an increase in the total muscle cross-sectional area. Whether there will also be hyperplasia (an increase in the number of muscle fibres) is more questionable, unless there is an injury of the muscle tissue. Besides an increase in cross-sectional area, there may be other structural changes of importance for force development such as increases in: (i) the length of muscle fibres due to an increase in the number of sarcomeres; (ii) the density of contractile proteins; and (iii) the intra- and extracellular connective tissue matrix, enhancing the proportion of the sarcomere force that can be transmitted to the skeletal system (Enoka 1988).

**Muscle training methods**

In principle, resistance exercise can be performed with static (isometric) or dynamic muscle actions. Dynamic exercise can be performed at a constant velocity throughout the range of motion, isokinetic action, or with a constant or variable load, but without control of movement velocity. The activity can be concentric or eccentric. In addition,
Electrical muscle stimulation can be used. The advantages of static exercise are that it can be used in the presence of limited joint mobility and also that a pain-free joint angle can be chosen. However, there will be no training through the full range of motion and static exercises are often considered ‘non-functional’.

Dynamic exercise with a constant weight can be easy to arrange, but will not match the pattern of maximal torque through the range of motion and may not accommodate resistance to pain. Eccentric exercise, if not wanted, cannot easily be avoided. During dynamic exercise with variable resistance, the load and muscle torque can be matched and it may, thus, be very efficient. However, special considerations may be needed to accommodate the resistance to pain as, for example, only using part of the range of motion.

Isokinetic exercise—concentric or eccentric—is sometimes considered non-functional, but is a well-controlled method of activating the muscle through the full range of motion, reducing the risk for overload. However, at the end of the movement there will no load and thus no training effect, for example, for end of knee extension.

Eccentric muscle action results in higher maximal torque than concentric muscle action (Westing et al. 1988). Neural activation is, however, not larger with eccentric muscle action as indicated by maximal electromyographic (EMG) activity, which may be lower than in concentric actions. Eccentric muscle actions increase the risk for muscle soreness, as demonstrated in a number of studies (Fridén et al. 1983). However, by producing greater muscle tension, eccentric muscle actions may enhance protein synthesis in muscles and contractile tissues. The soreness after eccentric exercise will lessen over the first week of continuous training (Fridén et al. 1983). As eccentric muscle action is part of many natural movement patterns, especially when the stretch–shortening cycle is used, as in running, jumping or throwing, they should be included in training programmes at some point.

It has been demonstrated that a short training period before immobilization may help to prevent muscle hypotrophy and loss of strength (Appell 1990). This might be explained by a maintained effect on the adaptation of the muscle and may have clinical implications for elective surgical procedures. However, the clinical relevance and usefulness has to be studied further.

Electrical stimulation

Electrical muscle stimulation in the early phase of strength training or during immobilization may promote neural activation by the afferent signals, reducing any pain and preventing or overriding reflex inhibition. However, in order to enhance the central neural drive to the muscle, electrical muscle stimulation has to be combined with voluntary muscle activation. At voluntary activation the low-threshold motor units innervating slow-twitch type I muscle fibres will be activated first, followed by the high-threshold fast-twitch type II fibres at a higher stimulus intensity (Gollnick et al. 1974). During direct nerve stimulation, however, larger axons belonging to the fast-twitch type II fibres will be activated first, as they have the lowest excitability. Combining electrical nerve stimulation and voluntary muscle activation may have a synergistic effect. However, no convincing evidence is available of the recruitment order of motor units, and in fact during surface stimulation no difference from pure voluntary activation has been seen (Knaflitz et al. 1990).

The order of motor unit activation with electrical muscle stimulation depends on at least three factors: (i) the diameter of the motor axon; (ii) the distance between the axon and the active electrode; and (iii) the effect of input to motor neurones from cutaneous afferents by the electrical muscle stimulation. It has been demonstrated in a number of studies that a combination of electrical muscle stimulation and voluntary muscle activation may limit strength reduction during immobilization and in the early postoperative phase (Arvidsson et al. 1986; DeLitto et al. 1988; Wigerstad-Lossing et al. 1988). It will also act to maintain muscle oxidative enzymatic capacity (Wigerstad-Lossing et al. 1988).
Gender differences

There are indications that the effect of neural adaptation in strength training would plateau somewhat earlier in women than in men (Häkkinen 1994). The effect on muscle volume is, however, more dominant in men than in women reflecting different hormonal influences. In women there may still be rather large interindividual differences in training-induced muscle hypertrophy, depending on the actual serum testosterone level (Häkkinen 1994). Even if certain specific aspects may be taken into consideration when training men and women, the general principles and the phases during rehabilitation are the same, with similar recommendations.

Strength training for children and adolescents

No convincing scientific evidence supports beneficial effects of using strength training for prepubescent children. Until more convincing studies are presented strength training should be avoided in this age group or used with low intensity. Reasons for this are the imbalances between the maturation of the muscle, tendon and skeletal tissues. Using too heavy loads often results in frustrating set-backs due to overuse or overload pain symptoms. To this can be added the absence of strength gains compared with non-strength-training peers. Strength training can slowly be added to postpubescent children but the training should be carefully monitored to avoid excessive overload or overuse.

Cardiovascular endurance

This section deals with endurance for more prolonged exercise, which is dependent on the function of the cardiovascular and respiratory systems. The overall capacity of the system can be defined as the aerobic capacity (maximal oxygen uptake). Endurance is defined as the ability to sustain prolonged aerobic exercise with an acceptable level of homeostasis, expressed, for example, as a percentage of maximal oxygen uptake or absolute power output or distance with a specific speed. Another aspect is the economy for body movement, as for running, with less energy consumption at a more economic movement pattern.

There are many factors contributing to endurance for whole body exercise, including the capacity of the heart with respect to maximal cardiac output and stroke volume, the distribution of blood flow to active muscles and the ventilatory function and blood–gas exchange in the lung. Coupled to these is the metabolic capacity for aerobic metabolism in the muscles. It is now well established that the maximal perfusion capacity of particular skeletal muscles is well above that which can be used for exercise with larger muscle groups (whole body exercise) (Andersen & Saltin 1985) and that the regulation of the peripheral blood flow in relation to the pumping capacity of the heart is essential (Saltin et al. 1998). Capillary density, leg blood flow and maximal oxygen uptake all increase in similar proportion after physical conditioning (Saltin & Rowell 1980).

With immobilization there will be a rather rapid reduction in the activity of the mitochondrial enzymes in the muscles (Appell 1990), which would also have an impact on the regulation of the central circulation contributing to the increased heart rate at a certain submaximal exercise level. There will successively be a reduction in the circulatory capacity towards the untrained state. Studies of immobilization and training of one leg have demonstrated the interaction between peripheral and central adaptation to exercise at immobilization and training (Saltin 1986).

Training of cardiovascular endurance after sports injury

Endurance training should include activities that are easy to perform despite the presence of any local body restrictions, such as remaining pain, limited joint motion, reduced muscle strength and local muscular fatigability. By tradition, exercises where the effect of bodyweight is reduced, such as swimming, pool exercises, bicycling and rowing machines, are used.
An important consideration is the possible interaction of concurrent strength and endurance training (Leveritt et al. 1999). During such a training regime the increase in strength may be reduced or limited compared with pure strength training, but the understanding of such an inhibition is limited at present. Various combinations of resistance and endurance training have to be studied to further understand the interactions. Whether, and to what extent, this also applies to training starting from an immobilized state, as after an injury or surgery, is not well known. Hypotheses based on acute (caused by residual fatigue after endurance training) and chronic (lack of possibility to adapt metabolically and morphologically to both strength and endurance training simultaneously) effects have been suggested. Endurance training even with resistance training superimposed may tend to reduce or limit the increase in muscle fibre size seen after resistance training alone. These aspects have to be taken into consideration when designing training programmes, especially in the later stages after immobilization and sports injuries in athletes participating in sports that require both strength and endurance.

Principal rehabilitation techniques

A gradually increasing training programme, with appropriate and continuously adjusted exercises, is probably the most important tool during rehabilitation after a sports injury. The rehabilitation programme can be divided into three different phases, not rigidly separated from each other, but with a more or less obvious overlap of different treatment approaches between the different phases.

First phase (acute phase)

Rehabilitation goals during the first phase

- Minimize the magnitude of the injury by correct acute management.
- Reduce or minimize loading of the injured tissue.
- Increase the circulation of blood and joint fluid with low loading exercises performed several times daily to aid in the healing process, to preserve range of motion and to accustom the tissues to an increasing load.
- Maintain strength, endurance, range of motion and function for the rest of the body.
To achieve a more than voluntary activated muscle force, superimposed electrical stimulation can be used during early immobilization. For optimal muscle effect the electrical stimulation should be given simultaneously with maximal voluntary activation, as discussed above. No randomized and double-blind studies are found in the literature studying the effects of electrical muscle stimulation during sports injury rehabilitation. However, superimposed electrical stimulation has been suggested for the recovery of quadriceps and hamstring muscle strength during rehabilitation after reconstruction of the anterior cruciate ligament (Arvidsson et al. 1986; Wigerstad-Lossing et al. 1988; Snyder-Mackler et al. 1995). It should be recognized that the placebo effect may have contributed to these effects. Also, the positive effect of increased motivation, with increased training effort and compliance, may have contributed to the results.

**Second phase (training phase)**

*Rehabilitation goals during the second phase*

- Increase load tolerance of injured tissues.
- Improve balance and coordination.
- Improve muscle strength and power and muscular endurance by gradually increasing loads and degree of difficulty regarding balance and coordination, using more demanding sport-specific exercises.
- A gradual change from one to two training sessions per day at the beginning of this phase to three or four highly demanding sessions per week at the end of the phase.

*Rehabilitation techniques*

In the second phase the injured tissue and surrounding muscles can withstand higher loads and thus be more actively mobilized. The treatment programme should be concentrated on the recovery of muscle mass. The number of repetitions used in the various exercises during the first phase is therefore gradually lowered to 8–12. Two to three sets per exercise and three training sessions per week are recommended. At the same time the load is gradually increased with a goal to reach 70–80% of maximal. At this level an optimal muscle hypertrophy can be anticipated. No major inhibitory influences should be present and an adequate motor recruitment is necessary. Therefore, at this stage, the training can be non-specific and various weight-training equipment can be used, such as leg-extension, leg-curl and leg-press weight machines. The use of repetition maximum (RM, the maximal load that can be lifted a given number of repetitions within one set) is a common method of determining the load used in resistance training. For increased maximal strength 1–5 RM are used, whereas 8–12 RM results in muscle hypertrophy.

During the late phases of anterior cruciate ligament (ACL) rehabilitation, slow and moderate improvements of quadriceps muscle strength and size have been reported (Risberg et al. 1999). This may indicate that rehabilitation programmes should have a greater emphasis on training for muscle volume and strength (i.e. high-intensity weight training using heavy loads which activate both high- and low-threshold motor units). Restoring muscle volume and strength is a cornerstone in rehabilitation, therefore weight training, with the optimal amount of resistance required for a maximal muscle growth process to occur, must be considered an essential training programme variable.

Interestingly, in sports training, as opposed to the rehabilitation field, a number of empirically based strategies have been developed through the years to maintain a positive response to long-term weight training (Fleck & Kraemer 1997). Thus, in advanced weight training various systems (e.g. periodization or organization of training into distinct periods) and methods (e.g. supersets, forced repetitions, power factor training and pre-exhaustion) are used to avoid performance to plateau, and for bringing about optimal gains of strength and muscle hypertrophy (Sisco & Little 1997; Bompa & Cornacchia 1998). However, current scientific knowledge of these practical strength- and power-training principles is poor.

Isokinetic training and testing equipment has been recommended for many years. However,
the validity, reliability and sensitivity has been strongly questioned during recent years (Gleeson & Mercer 1996; Stone et al. 2000). In well-controlled studies, isokinetic training, previously thought to be superior, has been found to be inferior to other training modes (Stone et al. 2000). Isokinetic testing is still the most commonly used strength testing procedure, but it is suggested that greater familiarization with the technique is needed, and several pre- and postevaluations, in order to improve testing reliability and sensitivity (Gleeson & Mercer 1996; Stone et al. 2000).

**Closed and open kinetic chain exercises**

There is a debate concerning the use of closed versus open kinetic chain lower limb exercises during rehabilitation. Studies have compared the effect of closed versus open kinetic chain lower limb exercises on knee ligament strain (Henning et al. 1985; Yack et al. 1993), anteroposterior tibiofemoral translation (Panariello et al. 1994) and patellofemoral compression forces (Gooch et al. 1993; Steinkamp et al. 1993). The importance of using closed kinetic chain rehabilitation (De Carlo et al. 1992; Shelbourne et al. 1995) and evaluation (Wilk et al. 1994; Greenberger & Paterno 1995) has been stressed. However, few studies have investigated whether the effect of closed versus open kinetic chain weight training on strength and performance differs.

Despite several studies (Yack et al. 1993; Beynnon et al. 1997) concerning safety issues of closed and open kinetic chain exercises in ACL rehabilitation, and although some authors have advocated the sole use of closed kinetic chain exercises (Shelbourne & Nitz 1992; Bynum et al. 1995), it is concluded that both types of exercises can be performed in ways that do not place excessive strain on the ACL (Fitzgerald 1997).

Escamilla et al. (1998), comparing knee joint biomechanics while performing closed and open kinetic chain weight training at a 12 RM load, reported that peak ACL tension forces in open kinetic chain exercise were only 0.2 times body-weight, and non-existent in closed kinetic chain exercise. Factors such as joint compressive forces (e.g. axial loading) and joint geometry probably play integral roles in knee joint stability during closed kinetic chain exercise (Isear et al. 1997). Significant coactivation of the antagonists during maximal knee flexion/extension, indicating an inhibitory mechanism which prevents overloading of the joint and contributes to joint stabilization (Kellis & Baltzopoulos 1998), would explain the low ACL tension forces during open kinetic chain exercise.

**Free weights and weight machines**

Most athletes use resistance training, involving both free weights and weight machines, to improve strength and power. The pros and cons of training with free weights versus weight machines are, as closed and open kinetic chain training, widely discussed among athletes and coaches, among physical therapists, as well as in sports science. Differences of opinion exist as to which method results in optimal performance gains. Proponents of free weights emphasize functionality and a direct application to sporting activities (Panariello et al. 1994). Conversely, the advocates of weight machines stress safety and less requirements of coordination as compared with free weight training (Kraemer & Fleck 1993).

In studies comparing free weights versus weight machines or closed versus open kinetic chain weight training, a dilemma exists in creating matching training conditions, e.g. equating total volume of training, total work and total training time. Moreover, as the weight does not vary in free weight exercise, the resultant torque, and therefore the required muscular contraction, does vary according to the mechanics of the specific exercise. Weight machines, on the other hand, operating through, for example, a cam, enable variable resistance throughout the range of motion of an exercise. This is an attempt to approximate the strength curve of the exercise, thus forcing the muscle to contract maximally throughout the range of movement.

A barbell squat is a free weight, closed kinetic chain exercise, involving muscles working across
multiple joints. Athletes using resistance training often include a barbell squat programme to improve lower extremity strength. Several studies have shown positive effects from a barbell squat exercise programme on strength and athletic performance (Thorstensson et al. 1976; Fry et al. 1991; Hickson et al. 1994).

Weight machine exercises, using muscles working across only single joints in an open kinetic chain, are also commonly used to improve lower extremity strength. The specificity of free weight and weight machine training is a critical issue, demanding accurate and sensitive tests to prove the superiority of one method over another. Dynamic (isotonic) testing such as a barbell squat or a vertical jump, activating the stretch–shortening cycle (Komi 1987), could be argued to be more valid than isokinetic testing when assessing sporting activities. However, the isokinetic knee extension exercise, though not activating the stretch–shortening cycle, has advantages such as greater control over velocity of motion, technique and extraneous movement, which facilitates measurement reliability and objectivity (Abernethy et al. 1995). Augustsson and Thomeé (2000) showed that a similar correlation existed for both closed ($r = 0.51$) and open ($r = 0.57$) chain testing with a vertical jump. However only 26–32% of jumping ability was explained by the strength in the closed ($r^2 = 0.26$) and open ($r^2 = 0.32$) chain tests (Fig. 12.3).

Much time and effort is initially spent learning proper technique when training with free weights. Conversely, weight machines are probably less difficult to master as they allow movement in only one plane and direction. Therefore,
it is theorized that weight machines cause greater gains of training load in short-term weight-training programmes. According to Sale (1988), increased performance as a result of weight training that lasts less than 20 weeks is associated mainly with neural adaptation, such as increased motor unit activity of prime mover muscles, and improved coordination, i.e. appropriate changes in the activation of synergists and antagonists.

Single-joint weight machine exercises allow a superior function of synergists and antagonists, enabling a high activation of motor units in prime mover muscles, as opposed to complex, multi-joint free weight exercises where full motor activation is probably more difficult to achieve due to greater demands of coordination. This theory is supported by recent studies in which open kinetic chain training resulted in larger increases of training load (Augustsson et al. 1998) and muscle hypertrophy (Chilibeck et al. 1998) than closed kinetic chain training in short-term weight training.

**Eccentric training**

Eccentric training is different from traditional concentric/eccentric training in that a higher load is possible during eccentric training. However, the same types of exercises, with the same number of repetitions, sets and frequency of training, can be used. The higher load can be accomplished in different ways. For example, in the sitting knee extension exercise the weight can be lifted concentrically with both legs and lowered eccentrically with one leg. The same principle applies to exercise in the standing position (Fig. 12.4). Another way is to have a training partner help during the concentric phase. Current isokinetic dynamometers allow specific heavy eccentric loading.

During normal concentric/eccentric weight training higher improvements have been noticed both in concentric and eccentric strength gains compared with concentric-only weight training (Dudley et al. 1991). Further, Higbie et al. (1996) reported higher eccentric strength gains using eccentric training and higher concentric strength gains using concentric training. A more interesting issue is, however, if eccentric strength training results in higher or faster strength gains, eccentrically or concentrically, than traditional concentric/eccentric training. No definite answer can be found in the literature, thus more research is needed. However, some studies, randomized and controlled, show better results with eccentric training programmes compared with concentric/eccentric training programmes for patients with chronic pain from the Achilles tendon (Mafi et al. 2000; Grävare Silbernagel et al. 2001).

**Muscle power training**

As muscle strength and load tolerance increases, the total load on the musculoskeletal system can be further increased. When close to normal strength is achieved, muscle power training and plyometrics (mainly reactive throwing and jumping exercises) can gradually be added to the rehabilitation programme, stressing time to peak muscualr tension as well as stressing the stretch–shortening cycle. It is recommended that maximal or close to maximal muscle power training and plyometrics (Fig. 12.5) are added to the training programme once or twice per week.
Three to five exercises with two to three sets of 10–15 repetitions and with a total training time of 10–15 min is sufficient. Thus the rehabilitation programme should use free weights (barbells and dumbbells), medicine balls, throwing implements and various jumping exercises. Häkkinen (1994) showed that explosive-type training (i.e. power training and plyometrics) had a positive effect on the maximal average integrated electromyography (IEMG)–time curve of the leg extensor muscles. Also, positive effects were shown with explosive-type training on the average isometric force–time curve with a higher speed of force development, compared with heavy resistance training.

Muscular endurance

If the general overall goal during the second phase is to also improve muscular endurance, many repetitions should be used with very short rest pauses of only a few seconds in between sets. This type of muscular endurance training can be performed daily with close to maximal intensity. Such a programme may increase endurance as well as strength and aerobic capacity of the muscle (Grimby et al. 1973).

Third phase (return to sports phase)

Rehabilitation goals during the third phase

- Regain maximal balance and coordination during maximally loading exercises as well as maximal endurance exercises.
- Regain maximal running, jumping and throwing capacity.
- Resume sports activity by implementing sport-specific exercises with maximal loading and maximal demand on range of motion, strength, endurance and coordination.

Rehabilitation techniques

The third phase can be arbitrarily defined to start when the muscle has regained sufficient volume, strength and endurance to approach a minimal level needed for various functional activities. Individuals have different functional needs, and the limitations will therefore vary according to the sport. A common feature, however, is the need for further improvement in muscle structure and activation, as well as improved motor control and coordination.
It has been noticed that despite ‘aggressive’ rehabilitation with early mobilization, full strength and muscle mass is not achieved 1 year after ACL surgery (Arangio et al. 1997; Pfeifer & Banzer 1999). Pfeifer and Banzer (1999) concluded, in their study of 39 patients with arthroscopically assisted ACL reconstruction, that insufficient rehabilitation schemes and not inhibition explained the large strength deficits seen 10–16 months after surgery. Despite this, many athletes resume their sports at 6 months and sometimes earlier after ACL reconstruction. It is thus not surprising that reinjury is common or that a new injury occurs. In many sports the musculoskeletal system is exposed to extreme loads. For example, during jumping the Achilles tendon is loaded with 5–10 times the bodyweight and the patellofemoral joint with up to 25 times the bodyweight.

There is, thus, a definite need for objective measurements of muscle function before full athletic activity is allowed. It is a future sports medicine challenge to develop a test battery of various strength tests, jumping, throwing and running tests, and agility and sport-specific tests. Tests need to be valid for the specific sport involved and tested for reproducibility.

During the third phase, training programmes to achieve and maintain gross motor functions must continue. More and more emphasis should be placed on specific training exercises for the specific sport that the athlete is to resume. Aspects of proprioceptive function, coordination and mechanical efficiency are discussed further in Chapter 13.

**Special aspects of cardiovascular endurance training programmes**

During rehabilitation after a sports injury it is important to try to maintain cardiovascular endurance. Thus regular bicycling, one-legged bicycling or arm cycling, an exercise programme in a pool using a wet vest or general major muscle exercise programmes with relatively high intensity and short rest periods (circuit weight training) can be of major importance. Increased maximal oxygen uptake has been shown to be maintained with a circuit weight programme (Gettman et al. 1979, 1982; Haennel et al. 1989). In these studies the participants trained 8–12 weeks, three times a week, using three circuits, with 8–10 exercises and with 20–30 s work and rest periods. Thus, an athlete sustaining an injury can maintain certain aspects of cardiovascular endurance by replacing, for example, running with an alternative form of cardiovascular endurance training. However, the peripheral muscular endurance close to the injury is more difficult to maintain and is limited to the recovery of the injury.

**References**


Chapter 13

Proprioception and Coordination

JUAN JOSÉ GONZÁLEZ ITURRI

Introduction

Proprioception can be defined as ‘a special type of sensitivity that informs about the sensations of the deep organs and of the relationship between muscles and joints’. Coordination can be defined as ‘the capacity to perform movements in a smooth, precise and controlled manner’.

Rehabilitation techniques increasingly refer to neuromuscular re-education. These methods have provided evidence of their efficacy and reliability in athletes, where just recovering the function of an injured joint is not sufficient, and perfect rehabilitation is required. Muscle-strengthening techniques are not sufficient since the neuromuscular and neuroarticular functions need to be re-educated. It is therefore time for rehabilitation to have a proprioceptive objective.

Ligaments, tendons and joint capsules do not just perform mechanical support functions or induce movement. They are also the bearers of deep sensitivity mechanoreceptors that gather information on the position of the joints, or how tendons and ligaments are stretched, and this information is sent to the base of the encephalon through the upper cord, initiating an effective response, resulting in dynamic stabilization and therefore a balanced reaction.

Neurological proprioception mechanisms

With regard to proprioception, as in many other fields associated with the nervous system, Sherrington’s contributions are fundamental due to his classification of sensitivity (Asirón Yribarren 1991):

1 Exteroceptive sensation. This is external surface sensation. It includes touch, pressure, pain, and hot and cold temperatures, which are globally known as general exteroceptive sensations; it also includes other more specialized sensations such as sight, hearing and smell, which together form special exteroceptive sensation.

2 Proprioceptive sensation. This covers a series of impressions on the functional state of the joints and muscles. If the individual is aware of these impressions, this is conscious proprioceptive sensation (awareness of passive and active movements, and the attitude or position of a somatic part of the body, kinaesthesia); if he or she is unaware of them, this is called unconscious proprioceptive sensation (relating to balance, tone and muscular coordination). There is also a special proprioceptive sensation, called labyrinthine.

3 Interoceptive sensation. This is found in the viscera and vessels (pain, oppression, etc.) and is of a general nature; it can be divided in areas such as gustatory and olfactory.

There are three types of sensation:

1 Deep sensation. This is the source of profound sensations, from pain to pressure, or sensations of movement.

2 Episensory superficial sensation. This corresponds to fine contacts, the special difference between sensations, the perception of temperature and the location of painful and temperature-related sensations.
3 Protopathic superficial sensation. This produces sensations of extreme pain and temperatures.

However, it is preferable to use less topographic and more functional classifications and descriptions. Gley (1914) talks of the sensation of attitudes and movements as a synonym for muscular sensations—awareness of the position of muscles in space and the sensation of passive movements, together with active movements and resistance. He establishes a series of divisions:

- Kinetic impressions, sensations of attitude, continuous, articular and muscular sensations.
- Kinetic impressions, sensation of passive movement. The perception of passive movements depends on articular sensitivity. Nevertheless, according to Gley, it is very difficult for joint sensations alone to determine the perception of passive movements.
- Kinetic impressions, sensation of active movements. Perception is a complex phenomenon combining tactile, joint and muscular impressions.

The afferent nerves responsible for kinetic impressions are cutaneous, articular and muscular:

- Cutaneous: these nerves are found in the fold and tension of the skin, and vary with the speed, energy and length of the contraction.
- Articular: articular anaesthesia reduces the perception of active and passive movements.
- Muscular: these nerves arise in the Golgi tendon organs and in the bones. The latter sense movement and the former sense tension (Grigg et al. 1982).

Feelings of central enervation, an efferent nerve, which comes from the brain and precedes the contraction. This is the sensation that an effort has to be made, of force released in the brain to provoke a movement.

Thounnard et al. (1986), in a critical study of the pathogenesis of a sprained ankle, reported that the passive structures are incapable of absorbing the energy produced when jumping on to an unstable surface. The time required to reach the calcaneotibial varus angle that causes the capsuloligamentous rupture is less than the reflex latencies observed in the muscles that stabilize the ankle, so it is assumed that only active muscular rigidity beforehand could protect this joint against a ligament injury.

For Strumpell (1924) the nerve endings in the skin are sensitive to stimulant disorders: mechanical, temperature-related and pain. He claims that the deep parts of the muscles, fasciae, periostea and joints can only feel mechanical pressure and distension stimuli, in particular tactile sensation referring to heat and cold, pain, pressure, mechanical distension, electrocutaneous sensation or sense of vibrations, and stereognostic sensation, which recognizes objects from their size, shape, hardness, etc.

For Sonjen (1986) the term exteroceptive is obsolete, whereas proprioceptive is used extensively. He points out the utility of classifying somatic sensation as superficial or deep, corresponding to the excitation of cutaneous sense organs and proprioceptors.

Sensitive receptor organs

Superficial somatic sensation is caused by exteroceptors, including: Pacinian corpuscles, mechanoreceptors for rapid, tactile and vibratory adaptations; Merkel’s disks, mechanoreceptors for slow adaptations, tactile and pressure sensitivity; Meissner corpuscles, also for rapid adaptations and tactile sensation; Krause bulbs or heat receptors; Ruffini corpuscles, slow adaptation mechanoreceptors for sustained pressure; and free endings for nociception and intense tactile and temperature stimuli (Burgess & Clark 1969; Burgess & Wei 1982).

Deep somatic sensation is due to proprioceptors, found in articular receptors like the Golgi organs, Ruffini corpuscles, Pacinian corpuscles, free endings and Krause’s end bulbs; and in muscular receptors like the laminated Pacinian corpuscles, Kühne’s neuromuscular spindles, Golgi–Mazzoni tendinous organs and free endings (Table 13.1).

Spinal level

As afferent elements in the spine, we have those already described, and as efferent elements, the
Table 13.1 Sensory receptor organs.

<table>
<thead>
<tr>
<th>Receptors</th>
<th>Location</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superficial somatic sensation (exteroceptive)</td>
<td>Pacinian corpuscles</td>
<td>Rapid adaptations, touch, vibration</td>
</tr>
<tr>
<td></td>
<td>Merkel’s disks</td>
<td>Slow adaptations, tactile sensation, pressure</td>
</tr>
<tr>
<td></td>
<td>Meissner corpuscles</td>
<td>Rapid adaptations, tactile sensation</td>
</tr>
<tr>
<td></td>
<td>Krause bulbs</td>
<td>Temperature receptors</td>
</tr>
<tr>
<td></td>
<td>Ruffini corpuscles</td>
<td>Rapid adaptations, sustained pressure</td>
</tr>
<tr>
<td></td>
<td>Free nerve endings</td>
<td>Nociception, tactile stimuli, endings</td>
</tr>
<tr>
<td>Profound somatic sensation (proprioceptors)</td>
<td>Golgi receptors</td>
<td>Deep sensation</td>
</tr>
<tr>
<td>Articular receptors</td>
<td>Ruffini corpuscles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pacinian corpuscles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Krause’s free endings</td>
<td></td>
</tr>
<tr>
<td>Muscular receptors</td>
<td>Pacinian corpuscles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kühne’s spindles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Golgi organs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Free nerve endings</td>
<td></td>
</tr>
</tbody>
</table>

final common pathway and as a response to the segmental and suprasegmental reflexes and voluntary impulses, the motoneuron axonal process located in the anterior horn of the spine contributing muscle fibres and provoking contraction.

Reflexes

In humans, the medullar reflex circuits play a decisive role in the control of movement. Some of them are monosynaptic, including the stretching reflex and passive muscular stretching, which excite the annulospiral endings.

When walking, monosynaptic stretching reflexes are inhibited. This has been verified by examining the behaviour of the potentials evoked when walking or standing; when walking, the group I afferent signals are blocked, both on the segmentary and supraspinal levels.

On the other hand, and unlike what happens when walking, segmentary stretching reflex activity helps to activate the muscles that stretch the legs during rapid movements such as running or jumping.

Polysynaptic reflexes are mutually inhibitory, since the same afferent nerves that determine the myotatic reflex provoke antagonist inhibition by means of an interneurone. Sherrington’s law of successive induction, a sequence of reciprocal inhibition, expresses the ease of contraction of a muscular group when antagonist contraction ceases.

The flexor withdrawal, or proprioceptive reflex, response to the nociceptive afferences on any structure, activating the flexor muscles and inhibiting the extensors, which provokes bending of the stimulated limb and the extension of the contralateral muscle.

Cerebral cortex

At the head of the hierarchy is the cerebral cortex, with two relief stations: the somatosensory area and the motor areas, through which a transcortical loop of variable gain and relatively slow execution is established, unlike the segmentary loops of less improvement but much faster execution.

For afferent nerves, there are two ways in which sensations reach the cerebral cortex to become conscious—they are thelemniscal system that carries precise time and space information, and the indirect extralemniscal system, with imprecise space and time characteristics.
For efferent nerves, there is the pyramidal path, a prolongation of the pyramidal neurones of the motor cortex that is directly and indirectly projected onto the motor neurones of the anterior horn of the medulla, to which the voluntary motorial function is attributed (Dietz 1987).

Finally, we have the extrapyramidal pathways that are projected on to the thalamus, nucleus basalis, red nucleus and pontine nuclei; they are involved in muscular tone, the regulation of movement and the regulation of posture.

Conclusions

Proprioception can be considered as the individual’s capacity to be aware of and recognize the instantaneous, static or dynamic situation of his/her body. According to classic proprioception classifications, it is responsible for capturing tension, force, spatial position, passive and active movements with speed, acceleration and amplitude, and the sensation of the structures, their situation and body limits (body schema).

All nerve endings, both superficial—general exteroceptive and special (sight and hearing)—and deep—general proprioceptive (articular and muscular), visceral, general interoceptive and special (taste and smell)—are involved in proprioception.

The nervous levels used include: the medullar segmentary level; the suprasegmentary level with the brainstem; and the cortical level, the activity of which is conditioned by specific and non-specific sensory afferent nerves, as well as the reticular thalamic system, or humoral regulation, cortical regulation or feedback.

Cutaneous, articular, tendinous and muscular afferent nerves follow the lemniscal path to the somatognosic areas, alert the cerebral cortex and produce proprioceptive sensation.

The use of proprioceptive mechanisms helps to correct or perfect motorial abilities (proprioceptive facilitation) and conscious proprioceptive perception provides awareness of posture and recognition of the body schema, allowing for their eventual correction.

Motor control and coordination

How the nervous system is organized to control and coordinate the neuromuscular function is much discussed and as yet unknown. It is an extraordinarily complex system, the capacity of which exceeds the most perfect of computers, and hypotheses on conscience, the thought process, motor planning and execution are impossible to prove. Such a complete organizational process involves structures such as: the cerebral cortex, which inhibits and excites the lower levels; the extrapyramidal system in charge of integrating the excitation and inhibition components, making them automatic by repetition; the corticospinal pyramidal path used to repeatedly excite the desired activity; and the basic spinal reflexes, which are the origin of all motor functions and sensitive feedback.

The human brain, unremarkable in appearance, and weighing approximately 1500 g, is evidently the most complex organization in the universe (Eccles 1969).

When referring to neuromuscular function, it is common to use the terms motor control, to define the excitation that voluntarily activates an isolated muscle, and motor coordination, a more complex neuromuscular activity that needs to integrate several elements to excite some muscles and inhibit others, to create patterns and sequences and thus to achieve functional body movements—something that is very important in sport.

The literature uses the terms motor control and coordination indistinctly, and includes many common models and theories concerning how it is organized and learned. In any circumstance, however, motor control is the result of a very important interaction between the individual, the activity and the surroundings (Fig. 13.1) (Acebes et al. 1996).

Measuring proprioceptive performance

Proprioceptive afferent signals take place and influence the motor response on several different
levels. Modulator influxes are added in the process of propagating and generating stimuli by means of the downward paths. Most of the interaction between motor and sensory responses takes place outside the voluntary perception level.

Psychophysical tests such as measuring the movement threshold or the capacity to reproduce an actively and passively preorientated angle are the test procedures used to assess proprioception. However, in sports our knowledge of proprioception is highly limited, for several reasons:

- It is not known how isolated proprioception mechanisms influence these tests. There are limited conclusions, for example, an improvement in a test result does not necessarily mean that there is a useful improvement in a person’s proprioceptive function.
- The tests are performed while resting. It is not possible to measure the proprioceptive function that the individual is capable of under effort.
- To what extent are the tests understandable outside voluntary perception?
- Angular reproduction tests are not constant for individuals.
- Training has a partially paradoxical impact. It can worsen test results or influence different psychophysical test procedures in opposite ways. Training can also influence the results of psychophysical test procedures.

Electromyography can provide an orientated analysis of muscular function as the end-point of regulatory proprioceptive function. Conclusions concerning possible influences can be obtained by observing muscular movement within the framework of the sensory motor function, but it does not measure an immediate proprioceptive function. However, it does have some advantages, because:

- Studies can be performed under conditions involving effort;
- The impact of training and the application of external stabilization aids, such as bandaging, can be studied by their effect on muscular action and regulation performance; and
- The pathological parameters of activation can be defined and identified.

Up to now, recognition procedure results have shown a decrease in the proprioception of major joints in relation to typical injuries and dysfunctions of the locomotive apparatus. It is difficult to evaluate the role of proprioception in the origin of injuries, since there are a lack of data concerning the proprioceptive situation before the injuries occurred.

Studying the use of proprioception in sport, it is possible to obtain epidemiological data and results from prior examinations:

- From the findings of psychophysical and electromyographic examinations, it can be deduced that injuries produce alterations to the special somatosensory and proprioceptive characteristics of performance.
- Before a major injury occurs, there are frequently earlier injuries that have been damaging the proprioceptive structures.
- Directed proprioceptive training essentially reduces the frequency of injuries.

**The effect of an injury on proprioception**

The natural history of, for example, the rupture of the anterior cruciate ligament of an athlete’s knee shows that, in spite of reconstructive surgical techniques and a good rehabilitation process
leading to a correct muscular situation, there is no guarantee that the treatment will be successful. This leads us to consider that there must be other factors, such as the enervation of the knee and the proprioceptive mechanisms, that are lost when the injury occurs (Table 13.2).

Capsular ligamentous elements, as primary passive structures, provide stability and normal kinematics to the knee joint. The presence of these receptor elements also means that, in addition to their mechanical function, they can act as dynamic organs by causing reflex synergic muscular activities.

The ligament injury, the resulting effusion, the immobilization required for treatment, and the following surgical trauma all produce alterations to proprioceptive sensitivity, creating confusing afferent messages that inhibit motor control and sensory motor programming, thus increasing the lack of stability by creating incorrect muscular responses. Knee stability cannot be restored without taking enervation into consideration (Jerosch & Prymka 1996; Lephart et al. 1997).

<table>
<thead>
<tr>
<th>Table 13.2 Knee joint receptors.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Receptors</strong></td>
</tr>
<tr>
<td>Ruffini corpuscles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Pacinian corpuscles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Golgi–Mazzoni corpuscles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Free nerve endings</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Proprioceptive re-education techniques

The treatments traditionally used for the recovery of injured athletes reprogramme the proprioceptive receptors and their afferent and efferent pathways. This is specifically the case for traditional kinesiotherapy, massage or modern isokinetic methods. They all stimulate the exteroceptive and proprioceptive systems.

In all these techniques, stretching, when correctly applied, has an important proprioceptive effect on the injured area, whether it be articular or muscular. Stretching not only gets the muscular receptors working, but also the articular level, improving integration of the body, which is fundamental in normal daily activities and sport (Commandre et al. 1996).

Proprioceptive techniques

Proprioceptive reprogramming can be achieved using the following techniques.
The use of unbalancing forces, using either unstable planes or applying destabilizers to an individual supported by a stable plane (Fig. 13.2). Both these kinds of imbalance are involved in sport, since the articular segment and the entire articular apparatus become involved in situations with components of either.

Reprogramming can also be achieved with the proprioceptive neuromuscular facilitation techniques proposed by Kabat at the beginning of the 1950s, consisting of the use of superficial and deep information, such as the position of joints, tendons and muscles, contact with the therapist’s hand, verbal orders and visual stimuli, all aimed at the work of the strongest muscle groups and leading to the functional improvement of other weaker, normally inhibited, groups.

Individuals make integrated movements that affect muscle groups that are used to working together. Complete kinetic chains work together following a series of rules: movement of the segments on the diagonal plane; a spiral component associated with each movement; and stretching stimuli which are used when applying maximum manual resistance.

For years, this method was essentially only used on patients with neurological disorders, but it is now frequently used in sports-relate pathologies.

Fig. 13.2 Different proprioceptive training situations. (a) Balancing on unstable planes; (b) balancing on one foot; (c) balancing whilst sitting; (d) sliding the foot along an unstable plane.
THE UPPER LIMBS

The proprioceptive re-education of a limb is always global, integrating the joints and the entire limb in the locomotive apparatus, and taking into consideration the limb’s function. Upper limbs are used for holding, picking up and throwing and lower limbs for walking, running, jumping and pushing oneself upwards. It is necessary to analyse what the segment has to achieve in each sport. Reprogramming a javelin thrower’s arm is not the same as reprogramming a swimmer’s arm (Fig. 13.3).

Exercises on unstable planes

Any of these exercises can be used, working in a closed kinetic chain. To increase the lack of stability, pressure is brought to bear on a Freeman board, and the exercise is performed in the quadruped position. Another variant is the use of a ball as a source of instability, while the individual is either standing or in a quadruped position. On occasion it is fundamental to throw the ball against a wall, forcing the individual to work to receive it.

Semiclosed and open kinetic chain exercises

Exercises using a stick held in both hands, while the physiotherapist unbalances it, are practical, very simple and can be graduated as required. This stimulates different positions, particularly in relation to the shoulder. This exercise can also be performed in a unilateral fashion.

The Kabat method

A basic exercise for the upper limbs is Kabat’s diagonal movements, using the scapulohumeral joint as a pivot and starting with a flexion and cubital inclination of the wrist, a pronation of the forearm, an extension of the elbow and with the shoulder positioned with extension, abduction and internal rotation. Applying manual resistance, the movement is towards a radial extension and inclination of the wrist, supination of the forearm, with the elbow continuing to be

![Fig. 13.3](a) Proprioceptive training of the shoulder. (b) Proprioceptive training of the shoulder by pushing against the physiotherapist.)
extended, and the shoulder moving towards flexion, adduction and internal rotation.

The second starting position can be a radial extension and inclination of the wrist, a pronation of the forearm, extension of the elbow, extension, adduction and internal rotation of the shoulder, moving towards a flexion and cubital inclination of the wrist, maintaining the elbow extended, and with the shoulder moving towards flexion, external rotation and adduction.

THE LOWER LIMBS

Using the Kabat technique with the upper limbs, two basic diagonal movements are used on which variations are then introduced. Something similar can be used after injuries to the lower limbs (Knoff & Voss 1968).

The first situation starts with a sole flexion and an eversion of the foot (with the knee bent or extended) with the hip extended with internal rotation, moving towards a dorsal flexion and inversion of the foot, and maintaining the knee in the original position, and the hip with flexion and external rotation (Blanc & Piur 1984; Bernier & Perrin 1998).

The second diagonal starts with a sole flexion and inversion of the foot, with the knee either bent or extended, and the hip extended with external rotation, moving towards an eversion and dorsal flexion of the foot, and flexion and internal rotation of the hip (Fig. 13.4).

Initially, the therapy starts with a perimaleolar massage, as there will probably be oedema, continuing with mobilization of the scar, if there is a scar, since scar tissue is occasionally adhered. Passive kinesiotherapy is then applied to all the ankle and foot joints, which helps to assess the state of the articular movements so that the athlete becomes aware of the state of his/her injured joint. These mobilizations should be smooth and never exceed the pain limit.

This is followed by active and selective unloaded work with the anterior tibial (flexion of the toes, varus of the foot and dorsal flexion of the ankle) and peroneals (plantar flexion of

![Fig. 13.4](a) General proprioceptive training acting on a lower limb with upper limb work. (b) Standing proprioceptive work.)
the ankle, flexion of the toes and varus of the foot). These exercises are performed both freely and with manual counter-resistance. The athlete should always perform the entire exercise in bare feet (Freeman 1965; Herveou & Messeau 1976; Herveou & Messeau 1981).

Stage 1

This exercise starts with learning crispation of the toes. The athlete should stand with the knees extended, trunk straight, feet parallel and in line with the coxofemoral joints and perpendicular to the frontal plane. The movement is towards a position in which the knees are extended and the trunk is straight and bent forwards, following a theoretical axis that passes through the malleoli. The purpose of this exercise is to work the musculature of the foot and adherence to the ground. It is the first level of proprioceptive information and the first state of alert.

Control of the anterior tibial is learned in the same way. The initial position is the same as for the previous exercise, but with a 30° flexion of the knees. The athlete bends forwards, imitating the initial stages of a fall, positioning the lower joints in an outwards turn. The physiotherapist should make sure that the contraction of the anterior tibial is intense and that the outer edge of the foot is supported (Fig. 13.5). The purpose of this exercise is the active maintenance of the internal longitudinal arch of the foot. It is the second level of proprioceptive information and the second state of alert.

Another exercise is with the two feet on the ground, knees extended, trunk straight, feet parallel and in line with the coxofemorals. The physiotherapist is positioned behind the athlete and on the side opposite the injury, pressing forwards and outwards.

Stage 2

What we have described above is the preparatory phase for the remaining exercises that work proprioceptively on the ankle, stimulating all the mechanoreceptors as much as possible. The exercises reproduce all the situations of the foot when it is in contact with the ground when walking, running and jumping. Therefore, the position of the foot during the exercises is very important, with unipodal support during the first half step, which is when the load is greatest.

The toes are crispated, the anterior tibial is tensed to maintain the internal longitudinal arch, the big toe is firmly positioned on the ground and the foot is perpendicular to the frontal plane. Contact with the ground is thus at the heel, the outer edge of the foot, the metatarsal heads, the tips of the toes and the entire plantar face of the big toe. Cutaneous information is maximal and reproduces the foot’s three support phases when walking. The knee is bent outwards at 30° to maximize the information captured by the proprioceptive mechanoreceptors, with the ligaments relaxed and maximum dynamic ankle stability.

![Fig. 13.5 Proprioceptive training on unstable planes with the physiotherapist’s help.](image)
Work with unipodal support starts at the second or third session of treatment. The athlete positions the injured limb in front of the healthy one, separated by a normal step length, reproducing the correction position of the foot (toes crispated, the anterior tibial tensed and the foot perpendicular to the frontal plane), with the knee bent to 30°, then returning to the initial position. The difficulty of the exercise can be increased by increasing the angle at which the knee is bent or extended.

**Stage 3**

At the third session, the athlete progressively starts the Freeman (BAPS) board exercises on the three axes of the rectangular plane. When the exercises on the rectangular plane are no longer difficult, work should start on the circular plane (Gagey 1993).

After the second week of therapy (after approximately 10 sessions of rehabilitation) the athlete should start jumping, first on the floor, then from the floor to the board and then from board to board, adding exercises involving catching a ball on the floor and on the Freeman board (Freeman 1965; Heurte et al. 1991).

At the last therapy sessions the athlete can run on a treadmill, beginning sports training combined with continuation of medical rehabilitation.

**Adapting proprioception to training**

Medicine’s role in sport is the treatment and, above all, prevention of athletic injuries. We must attempt, therefore, to adapt proprioception to training systems (Raybaud et al. 1988; Caraffa et al. 1996). Individual attention is not always possible or desirable for athletes; a structure accessible to several athletes at the same time needs to be considered. It is possible to combine sports training and proprioceptive training by installing an ‘athlete’s programme area’ next to the usual training area. These programme areas, which use proprioceptive re-education methods and techniques, can vary in infinite ways and be adapted to different sports.

Unbalanced surfaces can be grouped together, and when athletes have good stability on this equipment they can go on to other surfaces where the instability is greater. At the end of the progression, unipodal exercises can be performed with a ball, medicine ball and/or sandbag.

The material required is commonly found in rehabilitation centres:

- A rectangular board on two hemispheres, which will selectively work on the internal, external and anteroposterior muscular groups of the lower limb, depending on the position adopted.
- A circular board with a central spherical wedge, which works on proprioception in a broader sense. It regulates lack of balance, depending on the size of the hemisphere. It is even better if it is coated with a kind of synthetic lawn material, which adds cutaneous information to the proprioceptive reflex.
- A hanging balancing board, which provides maximum difficulty both for the lower limbs and for the trunk.

Within this framework, the athlete could work for 5 min on the balancing boards, going from the simplest to the most complex, ending with unipodal support and catching a ball. The training could continue with unipodal exercises on a soft carpet, where the athlete moves forward on the balancing bench to jump on a trampoline. From here he or she might move to the circumflex boards, after crossing a common area jumping from platform to platform.

**Coordination training**

Improving coordination depends on repeating the positions and movements associated with different sports and correct training. It has to begin with simple activities, performed slowly and perfectly executed, gradually increasing in speed and complexity. The technician should make sure that the athlete performs these movements unconsciously, until they finally become automatic.

In the first place, coordination training requires the athlete to be willing to undertake the activity
or to rest whenever he/she wishes; continuous perception information of the activity is needed, with perfect peripheral sensitive receptors from the subcortical centres. This creation of an automatic movement can only be achieved by voluntary repetition.

The basis of coordination training is the inhibition of the motor neurones that are not involved in the movement desired. Although direct training is not possible, this can be maintained by precisely executing certain activities. For this training, the individual has to be old enough to understand and follow instructions, be capable of learning and cooperating, and be capable of concentrating and avoiding fatigue. The receptors, paths and centres have to be intact and there must be a movement span of at least 30°, free of pain, at the joint affected by the muscle. Pain will lead to inhibition and, therefore, lack of coordination (Cerda et al. 1996).

Fig. 13.6 (a) Spinal column proprioceptive training whilst lying on one side. (b) Spinal column proprioceptive exercise whilst crawling.
Conclusions

For treatment to be effective it has to be applied early. The purpose of re-education, depending on the aetiology and severity of the injury, is to restore the athlete’s functional possibilities. It has to lead to a rapid return to physical activity compatible with the competition level.

Proprioceptive re-education has to get the muscular receptors working, in order to provide a rapid motor response (Scott et al. 2000). The treatment has to be adapted to each individual, considering the type of injury and the stress to which the athlete will be exposed when practising his or her sport.

There is no test for proprioceptive training. Its success can only be measured under controlled conditions in which the training is assessed and a reduction of the injury is confirmed (Quante & Hille 1999; Coarasa Lirón de Robles 2001).

Proprioceptive re-education is not a synonym of exercises on unstable planes. It re-educates ‘unstable’ athletes who are going to practise sport on unstable planes (Lorza 1991, 1998; Lephart et al. 1997; Guegan & Nicolas 1999; Romano 1999).

It is important to use proprioception when treating athletic injuries (Fig. 13.6). Global work has to be carried out on the injured limb, using the body’s own mechanisms, which are capable of being reprogrammed when practising sport again. It has to be carried out in an early stage and cannot guarantee that there will be no further relapses.

References


Chapter 14

Functional Rehabilitation and Return to Training and Competition

W. BEN KIBLER AND T. JEFF CHANDLER

Introduction

The goal of function-based rehabilitation programmes is the return of the athlete to optimum athletic function. Optimal athletic function is the result of physiological motor activations creating specific biomechanical motions and positions using intact anatomical structures to generate forces and actions. Sport-specific function occurs when the activations, motions and resultant forces are specific and efficient for the needs of that sport. We use the ‘critical point’ framework (Fig. 14.1) to understand sport specificity. Every sport has metabolic, physiological and biomechanical demands inherent in the way the sport is played, at whatever level of skill or intensity. Every athlete brings a specific musculoskeletal base to interact with the demands. Performance in the sport and injury risk are the results of the interaction. Sport-specific functional rehabilitation should focus on restoration of the injured athlete’s ability to have sport-specific physiology and biomechanics to interact optimally with the sport-specific demands.

Rehabilitation protocols are structured around frameworks of progressions of exercises, loads and applied and generated forces as the athlete improves his or her ability to generate and control sport-specific forces. Our rehabilitation protocol includes three phases: acute, recovery and functional (Fig. 14.2) (Kibler et al. 1992). The patient progresses through each phase in a ‘flow’ dependent upon reaching certain criteria of injury healing and achievement of control of certain physiological and biomechanical functions. The acute phase is a broad, relatively non-specific phase with goals of:

1. Protection or healing of injured tissues.
2. Improvement of general flexibility and strength.
3. Improvement in the distant links in the kinetic chains involved in the sport or activity.

Fig. 14.1 The ‘critical point’ framework showing the relationship between sports demands and the musculoskeletal base in determining performance and injury risk.

Fig. 14.2 Phases of the rehabilitation process. The start of rehabilitation includes all causes of injury, but the end focuses on return to function.
The recovery phase is a relatively long phase with goals of:
1 Preparation for maximum physical activity.
2 Improvement in kinetic chain linkage sequencing.
3 Normalization of any distant kinetic chain breakages such as inflexibility or strength imbalance.
4 Setting up the strength and power base for sport-specific progressions.

The functional phase is a focused phase, with the goals of sport-specific flexibility, strength, physiological activations, biomechanical motions and functional progressions.

The functional phases should simulate the actual activities required in the sport. Return to play criteria are based on successful completion of the functional phase goals. The functional phase marks a transition between rehabilitation and prehabilitation, which can be defined as conditioning strategies in the formerly injured athlete to prepare him or her for the stresses and demands inherent in the sport, and which must be faced in return to sport (Kibler & Chandler 1994). Prehabilitation focuses on sport-specific musculoskeletal areas that have been shown to be weak or susceptible to injury, or have actually been injured, in a specific sport (Chandler 1995). In some sports, the musculoskeletal system adapts to repetitive use by manifesting areas of muscle weakness or muscle tightness, or both. For example, long-distance runners frequently are found to have hamstring and gastrocnemius inflexibility, baseball and tennis players are found to have shoulder internal rotation deficits, and female soccer and basketball players have weakness of their glutei and hamstrings. These deficits are still present after injury, and may be even worse as a result of the injury or treatment. The purpose of the prospective programme is to condition those areas of high tensile load and high injury risk, or those that are the ‘weak link’ after injury, in a particular sport. Prehabilitation exercises consist of strength, power, range of motion and endurance exercises for the musculoskeletal areas and movements identified in a particular sport. This chapter will describe the activities that should take place in the functional phase.

The other two phases have been discussed in earlier chapters.

Upon entry into the functional phase, the athlete must have healing of the anatomical lesion, restoration of the distant links in the kinetic chain, general flexibility, endurance and strength, and be able to achieve the motions and positions necessary to play the sport or activity. The therapist and doctor must also know the demands of the sport or activity and the level of preparedness of the musculoskeletal base to work towards sport-specific progressions.

**Sport-specific demands**

There are many similarities in the demands sporting activities in general place on athletes. Most sports require some level of energy expenditure to generate the forces necessary to move the body, ball or other object toward a certain goal (Fig. 14.3). Most require some stability of one part of the body to achieve some movement of another part (Fig. 14.4). Most require some degree of flexibility of some body parts (Fig. 14.5). However, each sport has its own set of inherent demands, and these demands may change with the level of skill, intensity or frequency with which the sport is played. Therapists and doctors need to know these parameters in order to prehabilitate the athlete properly. Many different systems may be set up to profile sports and their demands. They may be based on anatomical requirements (range of motion required, strength required), metabolic requirements (aerobic or anaerobic activity) or amount of contact (collision, frequent contact or minimal contact).

We use an evaluation system that profiles the physiological requirements and the biomechanical kinetic chains necessary to play the sport (Kibler 1990). This method allows good discrimination between sports and generates enough information about the specific sport to plan an adequate prehabilitation programme that will be sufficiently specific for the sport.

The five physiological parameters we use include flexibility, strength, power (force × distance...
per unit of time), anaerobic and aerobic endurance. We grade each sport from 1 to 3 based on the parameter’s importance in the sport: 1, the parameter is important for general body fitness (e.g. aerobic endurance in golf); 2, the parameter is important in performance or injury risk reduction (e.g. flexibility in tennis or basketball); and 3, the parameter is essential for maximum performance (e.g. aerobic endurance in running, strength in weightlifting). Each sport will then have a profile that can be used to meet the demands. Table 14.1 lists some sports and their resulting profiles.

Similarly, each sport can be characterized by the kinetic chains necessary to perform the sport. Kinetic chains are patterned, coordinated sequences of body segment activation and motion.
Running has a slightly different kinetic chain than sprinting. Tennis has a different kinetic chain than baseball pitching. Kicking a soccer ball has a different kinetic chain than jumping after a basketball. Table 14.2 lists some of the most common sports and their predominant kinetic chain patterns.

When the physiological and biomechanical requirements of the sport or activity are categorized, exercises and progressions of load and intensity can be employed to prepare the body for the sport. A runner will need back and leg flexibility exercises, progressive aerobic training with longer duration of training bouts, and eccentric strength training of the legs to absorb the closed chain loads in the stance phase. An ice hockey player will need anaerobic training for short bursts of activity and power training in both the arms and legs. A baseball pitcher will need arm and back flexibility, lower body strength, power training to develop force, and eccentric training for the arm to control deceleration loads.

Functional musculoskeletal prehabilitation exercises can be strenuous as they replicate the motions, positions, forces, intensities and muscle activations inherent in the sport. The athlete’s body should have enough general athletic fitness to allow these vigorous activities without undue injury risk. The anatomical lesion should be healed. The physiological muscle activations, flexibility and endurance capabilities should be capable of withstanding the strength and power exercises.

The physician and therapist should monitor and test for these parameters as rehabilitation progresses and as the athlete goes into the prehabilitation protocols. For throwing athletes, the general criteria should include hip and trunk control over the planted leg, trunk rotational flexibility and strength, control of scapular elevation and protraction, functional glenohumeral internal rotation, and rotator cuff co-contraction strength (Kibler et al. 2000). For running athletes, general criteria should include gastrocnemius, quadriceps and hamstring flexibility, hip and trunk control over the planted leg, trunk extensor strength and aerobic capacity.
Athletes frequently have areas of inflexibility and weakness in other parts of the kinetic chain that may contribute to the injury or to decreased performance, and should be included in the prehabilitation exercises. The frequency of these distant alterations may be as high as 50–85% in association with injury (Ekstrand & Gilquist 1982; Kibler et al. 1991; Knapik et al. 1991; Warner et al. 1992; Burkhart et al. 2000).

**Specific examples—prehabilitation protocols**

The prehabilitation protocols should be based on progressions of flexibility, strength exercises, increasing loads and durations, and progress from closed chain to open chain configurations. This is illustrated by specific protocols for: (i) tennis/baseball, for overhead activities; (ii) running, for long distance and long duration; and (iii) soccer, for high-intensity power running and jumping.

**Tennis/baseball**

**PREHABILITATION PLAN**

The athlete should begin a shoulder flexibility and strengthening programme focusing on increasing internal rotation range of motion and external rotation strength. Flexibility exercises are performed after the throwing activity is complete for the day. Due to the high tensile loads developed in the shoulder external rotators and due to the inflexibility of the athlete in this area, a flexibility programme is warranted. For strengthening, the goals of the training programme are to increase the strength of the scapular stabilizers and shoulder external rotators while maintaining explosiveness in those muscles. Plyometric/agility drills are incorporated into the prehabilitation programme as they may reduce the risk of injuries related to these ballistic movements. In addition to the prehabilitation programme, it is assumed the athlete participates in a general resistance training programme two or three times per week.

At this stage of healing and rehabilitation, physical therapy modalities (heat, ice, ultrasound, stimulation) are usually not indicated. Very little tissue alteration is expected at this stage of healing that would be modified by these modalities. Some heat may be used to help in increasing tissue pliability before the exercise bout, and cold may be occasionally used if there is some swelling postexercise.

**WEEKS 1–2**

**Flexibility** (daily, after activity):
- Cross body stretch, scapula stabilized
- Internal rotation, scapula stabilized
- Triceps stretch
- Racquet stretch (Fig. 14.6)

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Reps/Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross body stretch</td>
<td>3 × 30 s</td>
</tr>
<tr>
<td>Internal rotation</td>
<td>3 × 30 s</td>
</tr>
<tr>
<td>Triceps stretch</td>
<td>3 × 30 s</td>
</tr>
<tr>
<td>Racquet stretch</td>
<td>3 × 30 s</td>
</tr>
</tbody>
</table>

**Strength** (3–5 × week):
- Shoulder external rotators/scapula retraction with tubing
- Rows with dumbbells/tubing, elbows down
- Reverse flies with dumbbells/tubing

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Reps/Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder external rotators/scapula</td>
<td>3 × 25 reps (moderate speed)</td>
</tr>
<tr>
<td>Rows with dumbbells/tubing</td>
<td>3 × 15 reps</td>
</tr>
<tr>
<td>Reverse flies with dumbbells/tubing</td>
<td>3 × 15 reps</td>
</tr>
</tbody>
</table>

**Plyometric/agility** (3–5 × week):
- Five dot drill (Fig. 14.7)
- Hexagon drill (Fig. 14.8)

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Reps/Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Five dot drill (Fig. 14.7)</td>
<td>2 × 20 foot contacts</td>
</tr>
<tr>
<td>Hexagon drill (Fig. 14.8)</td>
<td>3 × 15 s, double leg</td>
</tr>
</tbody>
</table>
**Line drill (Fig. 14.9)**
- 2 × 20 double leg foot contacts
- 2 × 20 single leg foot contacts

**Footwork ladder, bounding (Fig. 14.10)**
- 2 × 20 double leg foot contacts
- 2 × 20 single leg foot contacts

**Footwork ladder, quick feet**
- 4 × 20 low-intensity foot contacts

**Footwork sprints (carioca, shuffle, back pedal, cross-over step, skip, high knee, etc.)**
- 5–8 × 9 m

**WEEKS 3–4**

**Flexibility** (daily, after activity):
- Cross body stretch, scapula stabilized 2 × 30 s
- Internal rotation, scapula stabilized 2 × 30 s
- Triceps stretch 2 × 30 s
- Racquet stretch 2 × 30 s

**5-dot drill**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Both feet</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>1–2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1m</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.5m</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Do each sequence with 5 repetitions

**Racquet stretch**
Use racquet as rigid rod between both hands to rotate shoulder

**Fig. 14.6** The racquet stretch. Use the racquet as a lever to rotate the shoulder into internal and external rotation (figure seen from behind).

**Fig. 14.7** The five dot drill. Each sequence should be done in the order listed and completed five times.

**Fig. 14.8** The hexagon drill. Do in the sequence shown and repeat three times.

**Fig. 14.8** The hexagon drill.
**Strength (3–5 × week):**

- Shoulder external rotation/scapula retraction with tubing: 4 × 25 reps (high speed)
- Rows with dumbbells/tubing, elbows down: 4 × 25 reps (high speed)
- Shoulder internal rotators with tubing: 4 × 25 reps (high speed)
- Medicine ball forehand: 2 × 25 reps
- Medicine ball backhand: 2 × 25 reps
- Medicine ball serve: 2 × 25 reps

**Plyometric/agility:**

- Five dot drill: 2 × 20 foot contacts
- Hexagon drill: 4 × 10 s, single and double leg
- Line drill: 2 × 20 double leg foot contacts
- Footwork ladder, bounding: 2 × 20 double leg foot contacts
- Footwork ladder, quick feet: 2 × 20 single leg foot contacts
- Footwork sprints (carioca, shuffle, back pedal, cross-over step, skip, high knee, etc.): 10–15 × 4.5 m

**Weeks** 5–6

**Flexibility (daily, after activity):**

- Internal rotation, scapula stabilized: 2 × 30 s
- Racquet stretch: 2 × 30 s
**Strength (2–4 × week):**
- Shoulder external rotation/scapula retraction with tubing: 2 × 25 reps (high speed)
- Shoulder internal rotation with tubing: 2 × 25 reps (high speed)

**Plyometrics/agility:**
- Five dot drill: 2 × 20 foot contacts
- Hexagon drill: 5 × 10 s, single and double leg
- Line drill: 2 × 20 double leg foot contacts, 2 × 20 single leg foot contacts
- Footwork ladder, bounding: 2 × 20 double leg foot contacts, 2 × 20 single leg foot contacts
- Footwork ladder, quick feet: 3 × 20 low-intensity foot contacts
- Footwork sprints (carioca, shuffle, back pedal, cross-over step, skip, high knee, etc.): 10 × 4.5 m

**Running**

**PREHABILITATION PLAN**

It is suggested that the athlete should start prehabilitation on a more forgiving surface for training purposes whenever possible, such as grass or a rubber track. The main goal of the prehabilitation programme is to allow the athlete to progress at 10–25% increases in work load, which are usually safe progressions. Strengthening the ankle while trying to improve ankle range of motion is warranted. Flexibility of the hamstrings and gastrocnemius is stressed. Low-intensity plyometric and eccentric activities may help prepare the athlete to withstand the repetitive loads of running, as well as improve explosiveness for the kick at the end of the race.

**WEEKS 1–4**

The athlete is running 1.6–4.8 km, 3–4 days per week.

**Flexibility (daily, after activity):**
- Lying hamstring stretch: 3 × 30 s
- Standing quadriiceps stretch: 3 × 30 s
- Leg twists for iliotibial band: 3 × 30 s
- Wall calf stretch: 3 × 30 s
- Trunk sit and reach: 3 × 30 s

**Strength (4–5 × week):**
- Squats (light resistance): 3 × 20 reps
- Lunges with light dumbbells: 3 × 20 reps
- Weighted heel raise: 3 × 20 reps

**Plyometrics/agility (3 × week, non-running days on grass):**
- Five dot drill: 3 × 30 foot contacts
- Line drill: 3 × 30 double leg foot contacts
**WEEKS 5–8**

The athlete is running 4.8–8 km on Monday and Wednesday and 8 km on Friday and Saturday.

*Flexibility* (daily, after activity):
- Lying hamstring stretch: $3 \times 30$ s
- Standing quadriceps stretch: $3 \times 30$ s
- Leg twists: $3 \times 30$ s
- Wall calf stretch: $3 \times 30$ s
- Trunk sit and reach: $3 \times 30$ s

*Strength* (2–3 × week):
- Squats (light resistance): $3 \times 10$ reps
- Lunges with light dumbbells: $3 \times 10$ reps
- Weighted heel raise: $3 \times 10$ reps

*Plyometrics/agility* (non-running days on grass):
- Five dot drill: $2 \times 20$ foot contacts
- Line drill: $2 \times 20$ single leg foot contacts
- Bounding drill: $2 \times 10–15$ foot contacts

**WEEKS 9–10**

The athlete is running 8–16 km on Monday and Wednesday and 16–24 on Saturday.

*Flexibility* (daily, after activity):
- Lying hamstring stretch: $2 \times 30$ s
- Standing quadriceps stretch: $2 \times 30$ s
- Wall calf stretch: $2 \times 30$ s

*Strength* (2 × week):
- Squats (light resistance): $2 \times 10$ reps
- Lunges with light dumbbells: $2 \times 10$ reps
- Weighted heel raise: $2 \times 10$ reps

*Plyometrics/agility* (non-running days on grass):
- Fartlek progressions—alternating low-intensity/high-intensity running bouts (20 m jog → 10 m run → 10 m sprint)

**Soccer**

**Prehabilitation Plan**

Soccer requires aerobic endurance, anaerobic capacity and power. It is common that injuries may recur if prehabilitation is not complete, because of the varied and large loads inherent in the demands.
WEEKS 1–3

Flexibility (daily, after activity):
Lying hamstring stretch 2 × 30 s
Standing quadriceps stretch 2 × 30 s
Leg twists 2 × 30 s
Wall calf stretch 2 × 30 s
Trunk sit and reach 2 × 30 s

Strength (3 × week, bodyweight as resistance):
Squats 2 × 20 reps
Forward lunges 2 × 20 reps
Side step/cross step lunges 2 × 20 reps
Heel raise 2 × 20 reps

Plyometrics/agility (3 × week):
Five dot drill 2 × 20 foot contacts
Line drill 2 × 20 double leg foot contacts
Footwork ladder, bounding 2 × 20 double leg foot contacts
Footwork ladder, quick feet 3 × 30 low-intensity foot contacts

WEEKS 4–6

Flexibility (daily, after activity):
Lying hamstring stretch 2 × 30 s
Standing quadriceps stretch 2 × 30 s
Wall calf stretch 2 × 30 s
Standing knee to chest 2 × 30 s

Strength (medicine ball or dumbbells as resistance):
Squats 3 × 15 reps
Forward lunges 3 × 15 reps
Side step/cross step lunges 3 × 15 reps
Heel raises 3 × 15 reps

Plyometrics/agility (3 × week):
Five dot drill 2 × 30 foot contacts
Line drill 2 × 20 double leg foot contacts
Footwork ladder, bounding 2 × 20 single leg foot contacts
Footwork ladder, quick feet 4 × 30 low-intensity foot contacts
Footwork springs (carioca, shuffle, back pedal, cross-over step, skip, high knee, etc.) 5–8 × 18 m
Weeks 7–10

Flexibility (daily, after activity):
Lying hamstring stretch 2 × 30 s
Standing quadriceps stretch 2 × 30 s
Wall calf stretch 2 × 30 s
Standing knee to chest 2 × 30 s

Strength (3 × week, medicine ball or dumbbells as resistance):
Squats 2 × 10 reps
Forward lunges 2 × 10 reps
Side step/cross step lunges 2 × 10 reps
Heel raises 2 × 10 reps

Plyometrics/agility (3 × week):
Five dot drill 2 × 20 foot contacts
Line drill 2 × 20 double leg foot contacts
2 × 20 single leg foot contacts

Criteria for return to sports

The injured athlete must be anatomically healed and must complete all of the rehabilitation stages to be considered for return to sport. Understanding the inherent demands allows an objective set of criteria for return. Throwers must have sport-specific functional range of motion of arm joints, stable scapula, trunk core stability and plyometric power development in the leg, no matter where their original injury occurred. Runners must have leg flexibility, trunk core stability and aerobic endurance, no matter where their original injury occurred (Kibler et al. 2000). All injured athletes must then complete sport-specific functional progressions of throwing, running, jumping, kicking or swimming. These exercises demonstrate that the athlete has the physiology and biomechanics necessary to withstand the inherent demands of the sport.

Some authors have advocated return to sport based on quantitative criteria such as quadriceps/hamstring ratios in runners or internal rotation/external rotation ratios in throwers. These ratios are usually based on isokinetic data, which have not been conclusively shown to be associated with true muscle capability. These ratios can be helpful in determining weaknesses, but do not represent total capabilities, and should not be used as sole determinants of return to play. It is more important to know if the athlete can do all of the functions required in play, rather than if the athlete can generate torque at one joint.

It may appear that the criteria for return to sport are rigid or easily determined. Actually, the criteria should be employed with some leeway, based on the athlete and the situation. The athlete should not be allowed to ‘play yourself into shape’, with no guidelines or input. This allows deleterious kinetic chain substitutions and increases the chances of reinjury (Devlin 2000). However, the athlete should not be ‘rushed’ through the prehabilitation protocol, because the sport-specific exercises do impose significant demands, once again increasing the chances of reinjury.

Experience has taught us the value of gradual return to play, with emphasis on steady progression of exercises, strict attention to the preparedness of the musculoskeletal base for new exercises, involvement of the kinetic chain, and resolution of both the local and distant alterations that may exist in association with the injury (Kibler et al. 1991, 1992).

Return to play is completed by a sport-specific functional progression of doing the specific
activity. Table 14.3 illustrates a sport-specific throwing progression (Andrews & Wilk 1996), while Fig. 14.11 illustrates a sport-specific soccer progression (Kibler & Naessens 1996).

**Conclusions**

Optimal return to sport requires not only rehabilitation of the general athletic fitness parameters for sport, but also the specific parameters for fitness in a particular sport or sporting activity. Devising a sport-specific rehabilitation programme requires knowing the end-point—the set of anatomical, physiological and biomechanical parameters the athlete will need in order to meet the demands inherent in participation in that sport. Sport profiling is key to this knowledge. The functional phase of the rehabilitation protocol must emphasize acquisition of sport-specific parameters, and then shades into prehabilitation, a maintenance programme to continue to improve sport-specific fitness. Criteria for return to play must emphasize gradual return to sport-specific functional progressions.

<table>
<thead>
<tr>
<th>Table 14.3</th>
<th>Interval throwing programme starting off the mound.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1: fastball only</strong></td>
<td>Use interval throwing to 36 m phase as warm-up</td>
</tr>
<tr>
<td>Step 1: Interval throwing</td>
<td>All throwing off the mound should be done in the presence of your pitching coach to stress proper throwing mechanics</td>
</tr>
<tr>
<td>15 throws off mound 50%</td>
<td></td>
</tr>
<tr>
<td>Step 2: Interval throwing</td>
<td></td>
</tr>
<tr>
<td>30 throws off mound 50%</td>
<td></td>
</tr>
<tr>
<td>Step 3: Interval throwing</td>
<td></td>
</tr>
<tr>
<td>45 throws off mound 50%</td>
<td>Use speed gun to aid in effort control</td>
</tr>
<tr>
<td>Step 4: Interval throwing</td>
<td></td>
</tr>
<tr>
<td>60 throws off mound 50%</td>
<td></td>
</tr>
<tr>
<td>Step 5: Interval throwing</td>
<td></td>
</tr>
<tr>
<td>30 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>Step 6: 30 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>45 throws off mound 50%</td>
<td></td>
</tr>
<tr>
<td>Step 7: 45 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>15 throws off mound 50%</td>
<td></td>
</tr>
<tr>
<td>Step 8: 60 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td><strong>Stage 2: fastball only</strong></td>
<td></td>
</tr>
<tr>
<td>Step 9: 45 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>15 throws in batting practice</td>
<td></td>
</tr>
<tr>
<td>Step 10: 45 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>30 throws in batting practice</td>
<td></td>
</tr>
<tr>
<td>Step 11: 45 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>45 throws in batting practice</td>
<td></td>
</tr>
<tr>
<td><strong>Stage 3</strong></td>
<td></td>
</tr>
<tr>
<td>Step 12: 30 throws off mound 75% warm-up</td>
<td></td>
</tr>
<tr>
<td>15 throws off mound 50% breaking balls</td>
<td></td>
</tr>
<tr>
<td>45–60 throws in batting practice (fastball only)</td>
<td></td>
</tr>
<tr>
<td>Step 13: 30 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>30 breaking balls 75%</td>
<td></td>
</tr>
<tr>
<td>30 throws in batting practice</td>
<td></td>
</tr>
<tr>
<td>Step 14: 30 throws off mound 75%</td>
<td></td>
</tr>
<tr>
<td>60–90 throws in batting practice 25% breaking balls</td>
<td></td>
</tr>
<tr>
<td>Step 15: Simulated game: progressing by 15 throws per work-out</td>
<td></td>
</tr>
</tbody>
</table>
Soccer field run

1. Sprint
2. Jump and head a ball
3. Run around cones
4. Run backwards
5. Pick up soccer ball
6. Dribble around cones
7. Three running jumps
8. Plyometric jumps over lines
9. Sprint

Fig. 14.11 Return to soccer functional programme. The drill encompasses all of the soccer-specific demands.

References


Chapter 15
Orthoses in the Prevention and Rehabilitation of Injuries
WILLIAM MICHEO AND ALBERTO ESQUENAZI

Introduction
Advances in fabrication, increased availability and patient interest over the past few decades have contributed to the proliferation of orthotic use in sports and other areas of medicine. At the present time a more complete scientific understanding of the functional mechanism of orthoses and how the body responds to them is needed to further increase the efficacy of this intervention. The usefulness of bracing for the prevention or treatment of lower limb joint instability, particularly of the knee, has been controversial, with clinical and biomechanical data pointing in opposite directions (Cawley et al. 1991). In addition, a lot of the information available to the clinicians about orthotics and their use in sports medicine is provided by the manufacturers of the devices.

Advances in applied biomechanics, motion and gait analysis have helped considerably by allowing critical parameters to be accurately measured in the athlete, thus infusing objectivity in the process of orthotic prescription and the clinical management of injury (Esquenazi & Talaty 1996; Zheng & Barrentine 2000).

However, the question of exactly what is the best orthotic device for a given injury, and how a brace may provide improved function for an athlete participating in sport, is still not completely clear. Prescribed orthoses may impair some of the normal biomechanical functions that coexist in the injured limb or may force compensation in other anatomical sites. For example, a knee brace prescribed to enhance mediolateral joint stability may produce a compensatory overactivation of the hamstrings that increases the likelihood of tendonitis (Esquenazi & Talaty 1996). In addition, they may increase energy expenditure or raise intramuscular pressure in the anterior leg compartments of the athletes that use them (Hirshenbofen et al. 1991; Styf et al. 1992).

The process of orthotic prescription is still subjective and evolving; it is dependent on the clinician’s understanding of the injury, acquired skill and knowledge of the principles of physics and biomechanics. In addition, it requires a working knowledge of brace types, materials, functional design and the requirement of the sport, which in general are not part of routine clinical training.

Development of new materials and fabrication techniques in the last decade have greatly increased the range of applicability of orthoses to a wide variety of clinical situations in sports medicine. Due to this, braces are lighter, stronger, better fitting, more durable and perhaps more effective. These advances in brace construction allow clinicians to generate almost any combination of biomechanical features, that may be specific for a particular clinical problem and most beneficial for the athlete (Esquenazi & Talaty 1996).

Terminology and classification
The International Standards Organization of the International Society of Prosthetics and Orthotics has defined an orthosis as ‘an externally applied
device used to modify structural and functional characteristics of the neuromusculoskeletal system’. Orthoses are applied to the external surface of the body to achieve one or more of the following: to relieve pain, resist abnormal joint motion, reduce axial load, stabilize musculoskeletal segments, prevent or correct joint deformity and improve function while permitting joint(s) movement (Redford 2000).

Orthotics in the past have been classified into two major categories: static and dynamic. Static orthoses are rigid and give support without allowing movement. Dynamic orthoses allow a certain degree of movement and are fabricated with a combination of rigid materials and movable parts such as joint, cables, rubber bands or springs. Sometimes, dynamic orthoses are also termed functional orthoses (Redford 2000).

Depending on the clinical situation, the desired effect and the biomechanical characteristics, orthoses are considered corrective or accommodative. Corrective devices are used to improve the position of the joint by modifying the alignment of skeletal structures. Accommodative devices attempt to maintain a given position and prevent further joint deterioration.

Furthermore, orthoses for use in sports—particularly those for the knee and ankle—have been classified as either prophylactic, functional or rehabilitative depending on the clinical situation in which they are used (Wirth & DeLee 1990) (Table 15.1). Prophylactic bracing is intended to prevent or reduce the severity of injury to healthy joints. Ankle prophylactic bracing has been used in basketball, volleyball and soccer (Thacker et al. 1999).

Functional braces are designed to control normal joint motion as well as to attempt to resist abnormal joint rotation and translation (Wirth & DeLee 1990; Cawley et al. 1991). Most functional braces are intended to address knee or ankle joint instability caused by ligamentous injury. A common clinical situation in which a functional brace is used is in the patient with a non-operated anterior cruciate ligament (ACL) injury who desires to continue to participate in sports (Micheo et al. 1995).

Rehabilitation braces are primarily intended to reduce pain and protect the joint immediately after an injury or during the initial postoperative period. These braces attempt to control the joint range of motion and provide protection in case of inadvertent loading that may result in injury (Wirth & DeLee 1990). These types of orthoses, along with early functional rehabilitation, have been used for the management of the injured medial collateral ligament of the knee (Reider et al. 1993).

In 1975 orthotic terminology was formally adopted to allow a simplified, standard, clear communication language among different clinicians. The terminology adopted encourages anatomical level description starting with the most proximal joint, followed by the orthotic adjective, rather than eponyms or proper names. If the orthosis is to involve the foot, ankle and knee it is termed knee–ankle–foot orthosis (KAFO). The particular material used and the desired functional effect of the orthoses should also be included as part of the design (Goldber & Hsu 1997).

Biomechanics

The major joints of the body can be seen as semi-constrained motion units with feedback control that respond to internal and external forces and allow a stable, complex, three-dimensional joint motion in both the loaded and unloaded conditions. During dynamic activity, powerful opposing muscle groups (flexors and extensors) exert moments of force working synergistically with soft tissue restraints (ligaments, capsule and tendons) as well as the surface geometry of the bones to position the joint for optimal loading.
Orthotic systems usually function with a three-point force system in order to control joint movement (Buttler et al. 1983). Devices with a long lever arm and wide, close-fitting cuffs which distribute force over a large area have improved physical characteristics to control joint translation (Redford 2000). For example, a knee orthosis that extends up in the thigh and distally to the calf will restrain the flexion or extension motion based on its alignment. Rotational, translation and transverse forces are difficult to control with an orthotic system since orthoses achieve joint motion control through force transmission from the bone to the soft tissue and on to the brace structure, which allows brace migration and makes the system not rigid. These limitations in force control about a joint are particularly troublesome in patients with injury-related strength or proprioception loss (Janshen 2000).

Brace design should allow forces to be distributed over the largest area possible in order to reduce pressure concentration (force over area) and potential secondary soft tissue injuries (Buttler et al. 1983). Friction and shear stress should be controlled by appropriate construction and accommodating the soft tissue which interfaces between the brace and the joint it will act upon. fabrics. Plastics have the benefit of being lightweight, adjustable, cosmetic and having total contact in custom braces. Total contact construction is important to reduce an undesirable concentration of forces over soft tissues or pressure-intolerant areas such as bony prominences.

Composite materials and metal orthoses have the advantage of increased durability and, in the hands of a skilled orthotist, still maintain built-in adjustability. With the advent of new products such as carbon graphite and extruded plastic materials, we now have the additional advantage of maximum tension strength with lightweight design. The latter, when used for sports orthotics, results in a lighter, stronger and more durable device. The choice of materials in the fabrication of orthotics is expanding rapidly and a detailed review of them is beyond the scope of this chapter.

Orthotics can be custom fabricated or bought off the shelf, with off-the-shelf orthotics being the most commonly prescribed braces today. In some cases, because of the patient’s anatomy, difficulties with fit, size of the individual and joint incongruity, custom orthoses are required (Esquenazi & Talaty 1996).

When fabricating orthoses attention should be given to accommodating bony prominences, to alignment with the anatomical axis and to reducing ligamentous stress. The difficulty of application, maintaining adequate suspension and comfort are also of importance in achieving compliance with the use of these devices (Loke 2000).

A complete orthotic prescription should include the patient’s diagnosis, consider the type of footwear to be used, include the joints it encompasses and specify the desired biomechanical alignment, as well as the materials for fabrication. Communication with the orthotist, who will fabricate or fit the brace, is of utmost importance in order to obtain a good clinical result.

Clinical application

The successful use of an orthotic device depends on matching brace function with the patient’s functional needs. Appropriate orthotic application will result in restraint forces that oppose an undesired motion (Kilmartin & Wallace 1994). Often pathology of the muscle, tendon or ligament results in multiple deviations of the normal joint movement process. Primary causes are often masked by compensations and secondary deviations, which sometimes makes the clinical intervention difficult. Emphasis has been placed on initially identifying the primary causes of the particular disorder as the point from which to begin the clinical intervention (Esquenazi & Talaty 1996; Esquenazi et al. 1997).
Understanding clearly the pathology and the resultant compensations, the orthotic needs for a specific sport, how a brace may be used to meet these needs, and what effect the brace exerts on the overall performance of the mechanical system and the athlete are areas where research is in progress. The potential, protective and preventive use of orthoses in the athlete who has not been injured is controversial and further investigation is needed to clarify their role in this area.

Foot orthotics
Heel cord and/or plantar fascia tightness, calf weakness, increased pronation or a high arch are all conditions that alone or in combination can increase the likelihood of foot- and knee-related injuries in athletes (Fig. 15.1). At high risk for these injuries are athletes who participate in sports that require running or jumping (Frontera et al. 1994).

Optimizing the biomechanical posture of the foot structures, improving lower limb flexibility and strength, correcting or accommodating the presence of any abnormality in the architecture of the foot and improving shock absorption are all key interventions to prevent disabling injuries (Lee et al. 1987; Schwellnus et al. 1990; Redford 2000). Although statistics are not available, it is plausible that more prescriptions are written for foot orthotics than for all other orthoses combined. The use of orthoses has been advocated to improve foot posture and reduce the incidence of lower limb injuries during running or jumping (Fig. 15.2) (Orteza et al. 1992; Kilmartin & Wallace 1994). Limited information is available in the scientific literature to prove or refute the efficacy of foot orthotics for this type of activity.

Prefabricated orthoses (off the shelf) are less expensive devices that can be adjusted to provide adequate control or correction. In the hands of the experienced practitioner prefabricated orthotics are a very effective tool in the management of abnormal biomechanics of the foot. In some instances because of the wide variation in foot configuration, size and pathology, and when time and economic factors permit, custom orthotics should be considered as a better treatment option. As the name indicates custom orthotics are made specifically for an individual with their application.
particular functional needs in mind. On-site fabrication by a skilled, certified orthotist or other trained clinician is preferable to referral to an outside site.

When prescribing foot orthotics for sports, knowledge of the patient’s complaint, the pathology in question, functional anatomy, biomechanics, orthotic components, materials, sport requirements and, finally, recognition of the anticipated functional outcomes are essential for proper prescription (Orteza et al. 1992). Consideration should be given to the fact that external loads act as dynamic forces that may require different orthotic prescription and materials for different activities (Cornwall & McPoil 1999; Hertel et al. 2001; Subotnick 2001).

Foot orthoses are designed to be corrective or accommodative devices. Corrective devices that are meant to improve the position of the foot by correcting the alignment of skeletal structures can be manufactured from a variety of materials. Rigid materials such as thermoplastics, acrylic laminates and carbon graphite composites are used frequently for this (Fig. 15.3a). Cork and polyethylene are examples of materials used for semirigid orthotic devices (Hertel et al. 2001). The patient with a pronated flexible foot and metatarsal stress fracture may benefit from the use of a semirigid orthotic device as part of a complete rehabilitation programme (Chandler & Kibler 1993; Cornwall & McPoil 1999).

Accommodative devices made of soft, open or closed cell foams alone or in combination with semiflexible materials are meant to provide cushioning and support to an already deformed limb structure (Fig. 15.3b) (Subotnick 2001). They prevent further deformity while attempting to improve function and to control the direction and/or alignment of the joint movement. Individuals with rigid high arched feet and recurrent injury such as Achilles tendonitis or plantar fascitis are candidates for a soft orthotic device (Mascaro & Swanson 1994).

Appropriate posting, longitudinal or transverse arch build-ups, heel lifts, pressure relief areas or other special modifications, as well as application of different materials in the same orthotic device, can be integrated in the design and combined at the time of fabrication (Subotnick 2001). There are several factors to consider when selecting materials used in custom orthotic fabrication; these include the type of sport, the physical demands and the length of time the orthosis is to be used. For example, a triathlete or long-distance runner may need a device that is rustproof and designed to function with perspiration or wet skin. If the brace is intended for use for a few weeks a prefabricated device may suffice; however, if intended for longer term use (several months) then a more durable custom brace constructed out of carbon graphite may need to be considered.

The amount of corrective or controlling forces to be applied and the amount of axial loading, as well as the amount of cushioning that is desirable without introducing mediolateral ankle instability, should be considered (Schwellnus et al. 1990; Brown et al. 1996; Hertel et al. 2001). One example of this custom application is the use of TPE® (thermoplastic elastomer) for the longitudinal

![Fig. 15.3](a) A carbon graphite corrective foot orthosis, and (b) a moulded accommodative foot orthosis.)
arch, Pelite® as a moulded substrate and Poron® for shock absorption, all in the same foot orthotic.

In the very active athlete, accommodative orthoses may need to be replaced often to adjust for material deterioration, while in the child or teenager, replacement should be done to accommodate for growth. The cost of these custom-made devices, depending on regional variations and availability, has a wide spectrum. For the less complex accommodative devices price is not significantly different to that for prefabricated ones. In some areas, fabrication and delivery from the time of moulding can be accomplished in less than 48 h. This time may be further reduced with the use of computerized design and manufacturing systems (Fig. 15.4). These two factors—cost and fabrication time—need to be considered in the selection of custom orthotics.

The shoe is an integral part of any lower extremity orthotic that includes the foot, as it will serve as the foundation for the device and will directly impact on its function. The correct shoe size should be measured with the athlete standing after a period of exercise, since with weight-bearing and exercise the foot configuration will change. The shoe should be manufactured out of a breathable material, have a supportive counter and deep toe box with a wide heel and a flexible toe break (Fig. 15.5) (Janisse 1993; Johnson 2001). The users of fluid-filled heel chambers should keep in mind the possibility of increased ankle mediolateral instability. A removable insert with longitudinal arch support is a good option, and if a foot orthosis is to be used, the shoe insole should be removed.

When a custom orthosis is ready it should be evaluated on the patient to assure proper fit and function. When these characteristics are achieved, supervised wearing time should be gradually increased.

In prescribing orthotics for paediatric patients, such as in runners, tennis or soccer players, consideration should be given to the fact that throughout childhood growth and development
act as dynamic forces that require frequent orthotic re-evaluation and modification. Not infrequently, as the patient approaches adolescence, cosmetic considerations and peer pressure supersede functional considerations, which impacts on orthotic use, compliance and function.

Ankle orthotics

The ankle sprain is one of the most common injuries in athletes, particularly in sports where participants frequently jump and land on one foot or are expected to make sharp cutting manoeuvres, for example basketball, soccer, football and volleyball (Lee et al. 1987; Frontera et al. 1994). Because ankle sprains are common and may result in days or weeks lost from practice and competition, efforts have been made to prevent such injuries through directly protecting the athlete with better shoes, ankle wrapping, taping or bracing (Tropp et al. 1985; Thacker et al. 1999).

Ankle braces range from simple off-the-shelf models that provide light support to custom hinged models that help control significant ankle instability. The important function is to limit plantar flexion and inversion, which is the position of instability for the ankle. For example, an articulated ankle foot orthosis with plastic moulded side pieces allows dorsiflexion and plantar flexion but very little inversion or ever-sion (Wilkerson 1992).

Several interventions that could lower the rate of occurrence of ankle sprains in a variety of sports have undergone scientific review (Tropp et al. 1985). Clinical trials in soccer and volleyball players suggest that training agility, flexibility, specialized ankle disc training and education reduce the risk of ankle sprain (Thacker et al. 1999).

In the past, taping the ankle has been the preventive method of choice for coaches and trainers in many sports. Data from controlled trials indicate that taping can prevent ankle sprains, despite the fact that tape loosens in approximately 10 min and provides little or no measurable support to the inverting ankle within 30 min (Liu & Jason 1994; Manfroy et al. 1997). This residual protection may be associated with increased proprioception that allows the peroneal muscles to react more rapidly to inhibit extreme ankle inversion.

Clinical studies have not shown that the height of the shoe affects the incidence of ankle injury (Barrett et al. 1993). This, in addition to the high cost of taping to prevent ankle injury, has led to the widespread use of ankle orthoses in the treatment and prevention of ankle injury. Semirigid orthoses provide external support, may enhance proprioception and may be adjusted more easily than tape (Thonnard et al. 1996; Baier & Hopf 1998). Data from randomized controlled trials demonstrate the effectiveness of some of these devices, especially for the prevention of reinjuries, although clinical research indicates that some devices will be more effective or more acceptable to athletes than others (Thacker et al. 1999). For example, lace-up (Fig. 15.6a) or elastic ankle supports are inexpensive and reusable but may be uncomfortable and do not provide uniform compression. Stirrup-type orthoses have been effective in reducing the recurrence of ankle injury and are acceptable to wearers (Fig. 15.6b) (Surve et al. 1994). However, they may be expensive and because of difficulty with fit may decrease performance levels.

Based on the available medical literature as well as the clinical experience of the authors, athletes who suffer an ankle sprain should complete supervised rehabilitation before returning to practice or competition, and those athletes suffering a moderate or severe sprain should wear an appropriate orthosis for at least 6 months. Research suggests that the benefit of the orthosis may persist for up to 1 year after injury (Thacker et al. 1999). It is not clear that there is a benefit of using an orthosis if the athlete has not been injured.

Knee orthoses

In the autumn of 1984, the Sports Medicine Committee of the American Academy of Orthopaedic Surgeons conducted a symposium on knee bracing to classify the types of knee braces available and to review the existing research. The Sports
Medicine Committee developed three categories to describe the various types of knee braces available: prophylactic, rehabilitative and functional (Wirth & DeLee 1990; Cawley et al. 1991).

Prophylactic knee braces are those that are designed specifically to prevent or reduce the severity of injury to the knee resulting from an externally applied force. They are not designed to provide increased knee stability in the knee with a disrupted ligament. These braces are used almost exclusively in American football in an attempt to prevent ligament injuries. They are designed either with lateral or bilateral uprights, hinges and thigh as well as calf bands (Osternig & Robertson 1993; Munns 2000). Though prophylactic braces are used to protect the medial collateral ligament they may preferentially protect the ACL (Paulos et al. 1991). Prophylactic braces continue to be manufactured and used; however, their efficacy for use in Olympic sports has yet to be resolved in the minds of many clinicians. In addition, they may increase the risk of injury at other joints of the extremity such as the ankle.

Rehabilitation braces or postoperative knee braces are those braces designed to be applied immediately following injury or ligament reconstruction (Wirth & DeLee 1990; Munns 2000). These braces are characterized by greater length and usually incorporate a means for providing selective range of motion control by the use of hinges that can limit motion or be locked in one position (Reider et al. 1993). Typically, they consist of circumferential wraps and straps attached to medial and lateral side bars with a hinge attachment at the level of the knee joint line. If rotational control is desired they may require attachment to a footplate. These devices are commonly used and are widely accepted within the sports medicine community. A note of caution for the prescriber of the brace is that motion at the knee joint may exceed the limits set by the hinges by 10–20° and this needs to be considered in managing the patient.
Functional knee braces are those orthoses that are designed specifically to provide mechanical stability in knees with disrupted ligaments (Fig. 15.7). These braces have straps or cuffs and knee hinges, and function by limiting anterior tibial translation, knee hyperextension and, to a lesser degree, rotation. These functional knee braces are routinely used in the chronically unstable knee and most physicians recommend their use in this application (Bonamo et al. 1990; Gotlin & Huie 2000). While functional knee braces continue to be applied on ligament-deficient knees, they are also commonly used in knees with reconstructed ligaments. Thus, functional knee braces are applied as prophylactic devices to provide strain or overload protection to reconstructed, healing ligaments or grafts.

The preponderance of existing brace research addresses functional knee braces. It has attempted to measure tibial translation using a variety of methods including both manual displacements and knee arthrometer systems. Many published studies demonstrate that knee bracing does reduce anterior tibial translation in varying degrees under low, non-physiological loading conditions. Many authors have commented in the literature that functional knee braces do not provide increased mechanical stability under higher loads such as the physiological loads associated with sport (Cherf & Paulos 1990; Cawley et al. 1991).

It is interesting that in clinical and biomechanical studies that also conducted a subjective evaluation of the orthoses, over 90% of the patients reported significant functional improvement when using a functional knee brace (Cook et al. 1989; Cawley et al. 1991; Anderson et al. 1992). This is despite the fact that laboratory research clearly shows that the use of a functional knee brace may not protect the ligament at the physiological loads associated with sports, will have a significant cost in terms of energy expenditure and may lead to reduced performance (Highgenboten et al. 1991).

Cook et al. (1989) found that the application of a functional knee brace to an ACL-deficient knee enabled patients to generate torque and shear forces during a cutting manoeuvre that were near those generated by normal, non-ACL-deficient subjects and significantly better than cutting manoeuvers performed by unbraced ACL-deficient limbs (Cherf & Paulos 1990; Hertel et al. 2001).

The stabilizing effect of functional knee braces has been correlated with knee stability and muscle firing patterns (Anderson et al. 1992; Wojtys 1993). When anterior tibial translation loads were applied to the partially loaded limb, reductions in tibial translation of 15–30% with muscles relaxed and 60–80% with the muscles contracted were documented. In addition, a slowing of muscle reaction time or a latent period between the onset of the load and the firing of the muscles was reported (Cawley et al. 1991; Wojtys 1993). It is interesting that the braces, which best controlled tibial translation, also exhibited the longest latent period for muscle firing.

Many clinicians are in agreement that functional knee bracing serves an important role in the treatment of non-reconstructed knee ligament injuries (Bonamo et al. 1990; Micheo et al. 1995). In general, studies that performed subjective evaluation of functional knee braces found
a positive response in those patients wearing orthoses.

Proprioception and proprioceptive deficits resulting from disruption of the knee ligaments has been a hotly debated topic within the orthopaedic and sports medicine community since the discovery of mechanoreceptors within the substance of both the ACL and the medial collateral ligament of the knee as well as the joint capsule (McNair et al. 1996; Birmingham et al. 2000). While there may be an interrelationship between neurosensory and neuromuscular control systems, and proprioception may provide positional joint awareness, it is not clear what is the mechanism for this phenomenon in the injured athlete. In theory, functional knee bracing may enhance position sense until the intrinsic neurosensory mechanisms are re-established and provide overload protection to the healing tissues.

For conditions that involve the patellofemoral joint, such as runner’s knee, a pullover knee sleeve with various types of cut-outs that are designed to support or stabilize the patella can be helpful (Paluska & McKeag 1999). For patellar dislocation or subluxation, a patellar-stabilizing brace that has a cut-out sleeve and heavier buttressing is indicated (Cherf & Paulos 1990). In patients with patellar tendonitis or jumper’s knee, an infrapatellar counterforce orthosis may be used (Fig. 15.8). The theory behind the use of this type of counterforce bracing is that it inhibits full muscular contraction and decreases the force on the partially injured musculotendinous unit (Crossley et al. 2001).

A treatment modality for patellofemoral pain that has gained recent acceptance is patellar taping. This technique tries to correct patellar malalignment, including subluxation, tilt and rotation, permitting progression in a rehabilitation programme of quadriceps muscle strengthening. Some clinical studies have shown good results in pain reduction and quadriceps muscle recruitment; however, radiological studies show that although the technique appears to position the patella medially this position is not maintained after exercise (Larsen et al. 1995; Kowall et al. 1996).

Low back orthoses

Low back pain is very common with a high incidence reported in men and women. The incidence of spinal injuries related to sporting activities has been estimated at 10–15%. Sports in which back injuries are among the most frequent reported injuries include golf, gymnastics, football, rowing, wrestling, weightlifting and tennis (Tall & DeVault 1993). The prevalence, however, varies with the patient’s age, the sport played and with the position played. In gymnastics, the incidence of back injuries is estimated to be 11% (D’Hemecourt et al. 2000).

The most common injury site has been reported to be the soft tissues. However, depending on the biomechanical demands of the sport, discogenic or posterior element injuries have also been described. Sports that place repeated flexion and rotation demands on the spine may lead to disc-related symptoms, while those that place repeated hyperextension demands on the spine may lead to posterior element injury such as spondylolysis (Kaul & Herring 1994; Standaert et al. 2000).
Young patients usually present with injury to the posterior elements while older athletes present with injury to the disc (Blanda et al. 1993; Gerbino & Micheli 1996; Young et al. 1997).

Bracing has been routinely used for management of spondylolysis (Blanda et al. 1993; Stinson 1993; Kaul & Herring 1994). Micheli has recommended the use of a rigid polypropylene brace (modified Boston overlap brace) constructed in 0° of lumbar flexion, prescribed for 23 h per day for up to 6 months (Fig. 15.9) (Gerbino & Micheli 1996). In the case of an acute spondylolysis, or even in chronic cases, where immobilization was attempted in order to achieve healing, his protocol has recently been modified to weaning from the brace as early as 4 months if there is evidence of healing by computed tomography scanning (D’Hemecourt et al. 2000). The use of less rigid materials, in combination with the avoidance of repeated hyperextension until the patient is asymptomatic, has been advocated by others, including our own clinical group with good results (Standaert et al. 2000).

The use of bracing for the treatment of disco-genic symptoms is more controversial (Stillo et al. 1992; Young et al. 1997). A soft brace or corset may be used in the initial stages for patient comfort with correction of posture and reduction of gross motion as the mechanism of action of the orthosis. If the patient’s symptoms are severe and worsen with a flexed posture, the use of a brace in 5° of extension has been advocated by some clinicians (Gerbino & Micheli 1996).

Treatment of stable compression vertebral fractures routinely includes the use of extension bracing. Depending on the location of the fracture, a thoracolumbar–sacral (TLSO) or lumbar–sacral orthosis (LSO) in extension is used. A custom-made TLSO or a Jewett-type extension brace may be used from 4 to 12 weeks depending on bony healing (Nachemson 1987).

Upper extremity orthoses

Hand and upper extremity injuries are among the most common injuries sustained by athletes. Unfortunately, there is a tendency to minimize their severity as the hand does not bear weight and the injuries rarely render the athlete unable to compete (Alexy & De Carlo 1998). Wrist and hand injuries are fairly common and described to occur in 3–9% of all athletic injuries (Rettig & Patel 1995).

Injuries can be traumatic or related to overuse. The most common problems of the elbow include tendinitis (lateral, medial and posterior), nerve dysfunction and ligamentous injuries. Lateral epicondylitis or tennis elbow is the most common overuse problem about the elbow in athletes (Plancher et al. 1996a; Sánchez 2001). Tenosynovitis of the wrist is quite common in sports; de Quervain’s syndrome (tenosynovitis of the first dorsal compartment) and extensor carpi ulnaris tendinitis are commonly seen in racquet sports and weightlifting (Rettig & Patel 1995).

The scaphoid is the most commonly injured carpal bone, accounting for 70% of all carpal fractures (Alexy & De Carlo 1998). The injury results from a fall on the hand with the wrist dorsiflexed greater than 90° and may be seen in a variety of sports. Dislocation of the proximal interphalangeal joint is also a common traumatic injury of the hand.
Nerve injuries can also occur associated with overhead sports, for example volleyball. Ulnar neuropathy at the level of the elbow can be seen in baseball pitchers, and median neuropathy at the level of the carpal tunnel can occur in racquet sports such as tennis and racquetball (Plancher et al. 1996b).

Orthoses can be used to manage acute and overuse injuries of the wrist and hand. Stable scaphoid fractures can be managed with a short forearm to thumb spica cast followed by the use of a thumb spica splint. Some athletes can return to sports participation with the use of a playing cast or a padded thermoplastic splint if there is evidence of healing (Alexy & De Carlo 1998). Dislocations of the proximal interphalangeal joint can be treated with closed reduction and the use of a digital extension block splint in approximately 30° of flexion (Sailer & Lewis 1995). After 3 weeks the patient may progress to the use of buddy taping until the patient has pain-free full range of motion.

Tendonitis of the first dorsal compartment can be treated with the use of a thumb spica splint in combination with steroid injection and physical therapy. A short course of immobilization with a wrist–hand orthosis in 15–20° of wrist extension can be considered for other types of acute tendinosis of the wrist.

Counterforce bracing with the use of a forearm support band may be used for the treatment of lateral epicondylitis. The orthosis should be used just distal to the elbow to alter the tension forces on the wrist extensor muscles (Fig. 15.10) (Tropp et al. 1985; Plancher et al. 1996a). As previously stated, the counterforce brace inhibits full muscle contraction thus decreasing the forces applied to the partially injured musculotendinous unit. It has also been shown to decrease elbow angular acceleration and decrease electromyographic muscle activity of the forearm muscles (Sailer & Lewis 1995; Plancher et al. 1996a).

Ulnar neuropathy at the level of the elbow cubital tunnel can be treated with protection using an elbow pad to avoid direct pressure over the nerve, or in severe cases with an elbow orthosis made of rigid thermoplastic material in 45° of flexion. In the early stages of recovery the orthosis should be worn during most of the day and as the patient’s symptoms improve it should be used only at night (Plancher et al. 1996a). Median entrapment at the level of the wrist should be treated with a splint that immobilizes the wrist in 0–5° of wrist extension. The orthosis should be worn at night and during the day when performing provocative activities that increase symptoms.

**Future research**

Research is underway with a different focus, which considers the effect of the orthotic intervention on each segment of the whole body, as well as on the total body response. This will help to clarify the brace–body interaction. Questions are being addressed such as what is the best orthotic alignment (arc of motion allowed), what is the global effect of the brace on the body, how does the body response to a brace change over time, and how do subject perceptions correlate to more quantitative outcome measures? Research can play an important role in advancing the use of orthoses by: (i) isolating key facets of orthotic
function; (ii) the introduction and refinement of analytical methods that provide a more explicit and quantitative measure of the action of the brace and body response; and (iii) the application of these methods to a detailed analysis of the entire system response rather than just the local effect at the site of the intervention.

An area where research is currently being conducted is intersegmental dynamic analysis. Knowledge of the highly coupled nature of human body dynamics is important in understanding the total body movement pattern, how adaptations are made to this pattern, and the origins of its dysfunction (Esquenazi & Talaty 2000; Talaty et al. 2001). Gage (1991) states ‘It is time to start thinking of the balance between joints. Many muscles span more than one joint and if a surgeon performs a procedure at the knee, it will also have an effect on the ankle and the hip’. His experience, like that of many clinicians, has indicated an awareness of the coupled nature of body interactions. While it may be easier to make the physical connection in the case of a biarticulate muscle, this principle applies to all muscles. The same principles apply to orthotic intervention. Research in our laboratory attempts to understand orthotic function and effects on this level (Talaty et al. 2001). To this end, tools have been developed which allow decomposition of the net effects of muscle groups on each segment in the body and an understanding of how these component effects sum to produce the overall effect. The role of our work is to take a rigorous technique and apply it to produce results, which can be generalized to better understand the specific dysfunction and what ‘type’ of optimized orthotic device should be used.

In addition to this biomechanical and gait laboratory research, clinical trials to appropriately assess the role of orthotics in sports medicine need to be performed. In the future, combining laboratory and biomechanical data with the results of well-structured clinical studies will allow sports medicine practitioners to have better information in order to address the orthotic needs of their patients.

References


Crossley, K., Bennell, K., Green, S. et al. (2001) A systematic review of physical interventions for patellofemoral


basic helix–loop–helix family (bHLH) 45, 47
basketball 7, 291
basophils 15
behavioural factors, rehabilitation outcomes 172
behavioural responses to injury 167–9
β-agonists 200
β-endorphin 195, 198–9
betamethasone acetate 193
biofeedback 172, 173
biomechanics
Achilles tendon injuries 64, 71–2
ACL autografts 88–9
collagen fibril size distribution and 60–2
ligament/tendon 57–60
see also tensile strength, ligament/tendon orthoses 302–3
tissue injury 62–4
biopsychosocial model 161–4, 175, 176
blood pressure 124
mean 122, 123, 124
blood volume 119–20, 121, 122
B-lymphocytes 19–20
body mass, effects of detraining 128, 131–2
bone 99–106
fractures see fractures
ligament/tendon insertion 60, 63
remodelling 99–100, 102
structure 99–100
subchondral, drilling 110–11
woven 102
bone marrow stimulation procedures, cartilage injuries 110–11
bone morphogenetic protein (BMP) 101, 105
braces see orthoses
bradykinin 14
bupivacaine 193, 200
bursae, corticosteroid injections 193–4
C1 esterase inhibitor 21
calcium apatite deposition 68
callus 101, 102
calpains 38
capillaries, response to injury 22–3
capillarization, muscle 130, 153, 260
carbohydrate utilization 125
carbon graphite 303
cardiac dimensions 122–4
cardiac function 120–2
cardiac index 122, 123
cardiac output 121–2
cardiopulmonary reflex 123
cardiovascular adaptations
effects of detraining 117–25, 137, 154–5
reduced training and 136
cardiovascular endurance 263–4, 270
carisoprodol 197
carpal tunnel syndrome 312
cartilage, articular 106–12
effects of immobilization 73, 154
injuries
classification 109, 110
mechanism 108
treatment 109–12
types and healing 108–9
structure 106–8
celecoxib 188, 189, 190
central biasing theory 218
cerebral cortex 276–7
ceruloplasmin 24
cervical spondylosis 223
cervical traction 221–2
Chediak–Higashi syndrome 17, 18
children
orthotics 306–7
strength training 263
chondral injuries 108–9
classification 109, 110
flap injuries 108, 109
fractures 109
matrix and cell 108–9
chondrocytes 106
autologous transplantation 111–12
chondroitin sulphate 107
chronic injuries 4, 5–6, 146
cryotherapy 204
NSAIDs 189
see also overuse injuries
chronic overload syndromes 148
chylomicron-retinyl esters 127–8
circulation 122–4
citrate synthase 126, 132–3
clinical symptom complex 146–7
clinics, sports injuries 4
clonidine iontophoresis 215
coaulation cascade 18–19
cocaine 200
codeine 195, 197, 200
cognitive appraisals
models 161, 162
sports injury 166, 167
cognitive-behavioural interventions 172–3
cognitive factors
adherence to rehabilitation 168–9
rehabilitation outcomes 171
cold therapy see cryotherapy
collagen
articular cartilage 106, 107, 108
‘crimp’ 57, 58
cross-links
in ACL autografts 89–90
mechanical disruption 58
effects of heat 207
effects of stretching 240
ligament/tendon 56–7, 58
effects of immobilization 74–5
normal ACL and ACL autografts 79–82
synthesis/deposition
effects of immobilization 75, 77
in ligament/tendon repair 65, 67, 68–9, 70
physical modalities and 73
in tissue repair 27
type I 56, 65, 81, 100
type II 56, 65, 79–81, 89–90
type V 79–81
typing, ACL grafts 79–82, 89–90
collagenase 39–40
collagen fibrils
biomechanics 57, 58–9
in healing ligament/tendon 65, 67, 68
mechanisms of damage 58–9
remodelling, healing ligaments 77, 84, 85–90
size distribution
in ACL allografts 83–5
in ACL autografts 78–9, 80, 81–2, 85–90
in chronic Achilles tendonitis 68
decorin antisense therapy and 91–2
exercised ligaments 75–6
immobilized ligaments 74–5
injured ligament/tendon 61–2
mechanical properties and 60–2, 82, 84
reconstructed ligament/tendon 62
communication, patient–practitioner 174
complement system 20–2
concentric exercise 262
conduction, heat 206–7
confusion 165
connective tissue, effects of stretching 240
constipation, opioid-induced 197–8
continuous passive mobilization (CPM), in chondral injuries 111
contracture, after injury 153–4, 237–8
contrast baths 206
contusion 8
convection, heat 207
conversion, heat 207
<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>coordination</td>
<td>277</td>
</tr>
<tr>
<td>coping behaviours</td>
<td>167, 169</td>
</tr>
<tr>
<td>corticosteroids</td>
<td>191–4</td>
</tr>
<tr>
<td>history</td>
<td>191–2</td>
</tr>
<tr>
<td>injections</td>
<td>192, 193–4, 196, 201</td>
</tr>
<tr>
<td>iontophoresis</td>
<td>213</td>
</tr>
<tr>
<td>mechanism of action</td>
<td>23, 192</td>
</tr>
<tr>
<td>phonophoresis</td>
<td>213</td>
</tr>
<tr>
<td>routes of administration</td>
<td>192</td>
</tr>
<tr>
<td>side effects</td>
<td>194</td>
</tr>
<tr>
<td>systemic</td>
<td>193, 201</td>
</tr>
<tr>
<td>corticotrophin-releasing hormone (CRH)</td>
<td>24, 25</td>
</tr>
<tr>
<td>cortisol</td>
<td>24–5, 131, 136</td>
</tr>
<tr>
<td>ultrasonography and training</td>
<td>211</td>
</tr>
<tr>
<td>cortisone</td>
<td>192</td>
</tr>
<tr>
<td>counselling</td>
<td>174</td>
</tr>
<tr>
<td>C-reactive protein (CRP)</td>
<td>24</td>
</tr>
<tr>
<td>creatine kinase (CK)</td>
<td>47–8, 131</td>
</tr>
<tr>
<td>creep</td>
<td>59, 233, 234</td>
</tr>
<tr>
<td>critical point framework</td>
<td>288</td>
</tr>
<tr>
<td>cross-training</td>
<td>137, 138</td>
</tr>
<tr>
<td>in water</td>
<td>208</td>
</tr>
<tr>
<td>cross-transfer effect</td>
<td>137, 138, 261</td>
</tr>
<tr>
<td>cruciate ligaments see anterior cruciate ligament; posterior cruciate ligament</td>
<td></td>
</tr>
<tr>
<td>cryokinetcs</td>
<td>205</td>
</tr>
<tr>
<td>cryostretch</td>
<td>205</td>
</tr>
<tr>
<td>cryotherapy</td>
<td>204–6, 224</td>
</tr>
<tr>
<td>application techniques</td>
<td>204–6</td>
</tr>
<tr>
<td>mechanisms of action</td>
<td>23–4, 204</td>
</tr>
<tr>
<td>‘cutting cone’</td>
<td>101</td>
</tr>
<tr>
<td>cycling</td>
<td>291</td>
</tr>
<tr>
<td>cyclo-oxygenase (COX)</td>
<td>13</td>
</tr>
<tr>
<td>inhibitors</td>
<td>15, 190</td>
</tr>
<tr>
<td>type 2 (COX-2) inhibitors</td>
<td>14, 190</td>
</tr>
<tr>
<td>cytokines</td>
<td>10–13, 25, 27</td>
</tr>
<tr>
<td>in acupuncture</td>
<td>199</td>
</tr>
<tr>
<td>anti-inflammatory</td>
<td>10–11, 13</td>
</tr>
<tr>
<td>bone</td>
<td>100</td>
</tr>
<tr>
<td>in muscle injury</td>
<td>39, 41, 42–4</td>
</tr>
<tr>
<td>proinflammatory</td>
<td>10–11, 14</td>
</tr>
<tr>
<td>in tissue repair</td>
<td>26</td>
</tr>
<tr>
<td>dash pot</td>
<td>234</td>
</tr>
<tr>
<td>debridement, cartilage injuries</td>
<td>109–10</td>
</tr>
<tr>
<td>decorin antisense therapy</td>
<td>91–2, 93</td>
</tr>
<tr>
<td>deep heating (diathermy)</td>
<td>207, 210–13</td>
</tr>
<tr>
<td>deep sensation</td>
<td>274</td>
</tr>
<tr>
<td>delayed-onset muscle soreness (DOMS)</td>
<td>62–3, 262</td>
</tr>
<tr>
<td>cycleraphytherapy</td>
<td>206</td>
</tr>
<tr>
<td>flexibility training and</td>
<td>236</td>
</tr>
<tr>
<td>iontophoresis</td>
<td>215</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>190</td>
</tr>
<tr>
<td>TENS 219–20</td>
<td></td>
</tr>
<tr>
<td>ultrasound therapy</td>
<td>212</td>
</tr>
<tr>
<td>depression</td>
<td>165</td>
</tr>
<tr>
<td>detraining</td>
<td>117–38</td>
</tr>
<tr>
<td>avoiding/reducing effects</td>
<td>135–7</td>
</tr>
<tr>
<td>cardiorespiratory effects</td>
<td>117–25, 154–5</td>
</tr>
<tr>
<td>definition</td>
<td>117</td>
</tr>
<tr>
<td>endurance performance</td>
<td>134–5</td>
</tr>
<tr>
<td>metabolic effects</td>
<td>125–9</td>
</tr>
<tr>
<td>muscular effects</td>
<td>129–34</td>
</tr>
<tr>
<td>strength performance and</td>
<td>135</td>
</tr>
<tr>
<td>syndrome</td>
<td>123</td>
</tr>
<tr>
<td>dexamethasone iontophoresis</td>
<td>215</td>
</tr>
<tr>
<td>dextromethorphan</td>
<td>200</td>
</tr>
<tr>
<td>dextropropoxyphene</td>
<td>200</td>
</tr>
<tr>
<td>diamorphine</td>
<td>200</td>
</tr>
<tr>
<td>diathermy</td>
<td>207, 210–13</td>
</tr>
<tr>
<td>dihydrocodeine</td>
<td>195, 200</td>
</tr>
<tr>
<td>diphenoxylate</td>
<td>200</td>
</tr>
<tr>
<td>diphenoxylate</td>
<td>200</td>
</tr>
<tr>
<td>direct current (DC)</td>
<td>213, 215</td>
</tr>
<tr>
<td>doping</td>
<td>200–1</td>
</tr>
<tr>
<td>dynamization</td>
<td>102</td>
</tr>
<tr>
<td>eccentric exercise</td>
<td>62–3, 262, 268</td>
</tr>
<tr>
<td>eccentric strength, training cessation and 135</td>
<td></td>
</tr>
<tr>
<td>eicosanoids</td>
<td>189</td>
</tr>
<tr>
<td>elastic energy, storage</td>
<td>57, 236–7</td>
</tr>
<tr>
<td>elastic fibres (elastin)</td>
<td>56, 57</td>
</tr>
<tr>
<td>elasticity</td>
<td>233, 234</td>
</tr>
<tr>
<td>effects of stretching</td>
<td>240</td>
</tr>
<tr>
<td>elbow</td>
<td></td>
</tr>
<tr>
<td>injuries</td>
<td>311, 312</td>
</tr>
<tr>
<td>orthoses</td>
<td>312</td>
</tr>
<tr>
<td>tennis see lateral epicondylitis</td>
<td></td>
</tr>
<tr>
<td>electrical stimulation</td>
<td></td>
</tr>
<tr>
<td>muscle</td>
<td>215–18, 262, 265</td>
</tr>
<tr>
<td>nerves</td>
<td>214, 218–20, 262</td>
</tr>
<tr>
<td>to promote tissue healing</td>
<td>214</td>
</tr>
<tr>
<td>single twitch</td>
<td>260–1</td>
</tr>
<tr>
<td>with tetanic trains</td>
<td>261</td>
</tr>
<tr>
<td>electrical therapy</td>
<td>213–20, 224</td>
</tr>
<tr>
<td>contraindications</td>
<td>214</td>
</tr>
<tr>
<td>fracture healing and</td>
<td>105</td>
</tr>
<tr>
<td>indications</td>
<td>220, 221</td>
</tr>
<tr>
<td>soft tissue repair and</td>
<td>73</td>
</tr>
<tr>
<td>electric heating pads</td>
<td>208</td>
</tr>
<tr>
<td>electroacupuncture stimulation (EA)</td>
<td>198</td>
</tr>
<tr>
<td>electrocardiogram (ECG)</td>
<td>124</td>
</tr>
<tr>
<td>electromagnetic field therapy</td>
<td>73, 223–4</td>
</tr>
<tr>
<td>electromyography (EMG)</td>
<td>278</td>
</tr>
<tr>
<td>in flexibility testing</td>
<td>239</td>
</tr>
<tr>
<td>in flexibility training</td>
<td>239–40</td>
</tr>
<tr>
<td>immobilization-related changes</td>
<td>151–2</td>
</tr>
<tr>
<td>emotional factors, rehabilitation outcomes</td>
<td>171–2</td>
</tr>
<tr>
<td>emotional responses to injury</td>
<td>164–7</td>
</tr>
<tr>
<td>mediating/moderating factors</td>
<td>166–7</td>
</tr>
<tr>
<td>β-endorphin</td>
<td>195, 198–9</td>
</tr>
<tr>
<td>endorphins</td>
<td>218–19</td>
</tr>
<tr>
<td>endothelial cells</td>
<td>22, 27</td>
</tr>
<tr>
<td>endurance</td>
<td>258–70</td>
</tr>
<tr>
<td>cardiovascular</td>
<td>263–4, 270</td>
</tr>
<tr>
<td>definition</td>
<td>263</td>
</tr>
<tr>
<td>muscle</td>
<td>260, 269</td>
</tr>
<tr>
<td>performance, detraining effects</td>
<td>134–5, 138</td>
</tr>
<tr>
<td>training, after injury</td>
<td>263–4, 270</td>
</tr>
<tr>
<td>enkephalins</td>
<td>195, 198</td>
</tr>
<tr>
<td>enthesis</td>
<td>60, 63</td>
</tr>
<tr>
<td>eosinophils</td>
<td>17–18</td>
</tr>
<tr>
<td>epipipical superficial sensation</td>
<td>274</td>
</tr>
<tr>
<td>epidemiology of sports injuries</td>
<td>3–8, 145</td>
</tr>
<tr>
<td>in international competitions</td>
<td>7–8</td>
</tr>
<tr>
<td>patterns of injuries</td>
<td>4–7</td>
</tr>
<tr>
<td>epinephrine (adrenaline)</td>
<td>119–20</td>
</tr>
<tr>
<td>erythropoietin</td>
<td>25</td>
</tr>
<tr>
<td>ethyl chloride spray</td>
<td>205</td>
</tr>
<tr>
<td>etodolac</td>
<td>188</td>
</tr>
<tr>
<td>exercise</td>
<td></td>
</tr>
<tr>
<td>ligament tensile strength</td>
<td>75–8</td>
</tr>
<tr>
<td>muscle regeneration</td>
<td>51–2</td>
</tr>
<tr>
<td>see also mobilization; training</td>
<td></td>
</tr>
<tr>
<td>explosive-type training</td>
<td>269</td>
</tr>
<tr>
<td>extensibility</td>
<td></td>
</tr>
<tr>
<td>definition</td>
<td>233, 234</td>
</tr>
<tr>
<td>effects of warm-up</td>
<td>250</td>
</tr>
<tr>
<td>exteroceptive sensation</td>
<td>274, 275</td>
</tr>
<tr>
<td>extracellular matrix</td>
<td></td>
</tr>
<tr>
<td>articular cartilage</td>
<td>106–8</td>
</tr>
<tr>
<td>proteolysis</td>
<td>39–40</td>
</tr>
<tr>
<td>TGF-β actions</td>
<td>43</td>
</tr>
<tr>
<td>extrapyramidal system</td>
<td>277</td>
</tr>
<tr>
<td>factor XII</td>
<td>21</td>
</tr>
<tr>
<td>factor H</td>
<td>21</td>
</tr>
<tr>
<td>factor I</td>
<td>21</td>
</tr>
<tr>
<td>faradic stimulation (alternating current)</td>
<td>213, 215, 216</td>
</tr>
<tr>
<td>fat, body, effects of detraining</td>
<td>128, 131</td>
</tr>
<tr>
<td>fatigue, after injury</td>
<td>258</td>
</tr>
<tr>
<td>fatigue failure, Achilles tendon</td>
<td>72</td>
</tr>
<tr>
<td>fear</td>
<td>165</td>
</tr>
<tr>
<td>femoral-medial collateral</td>
<td></td>
</tr>
<tr>
<td>ligament-tibial (FMT) complex</td>
<td>77</td>
</tr>
<tr>
<td>fenoprofen</td>
<td>188</td>
</tr>
<tr>
<td>fentanyl patch</td>
<td>197</td>
</tr>
</tbody>
</table>
fever | 14, 24  
*fibrinogen* | 127  
*fibroblast growth factor (FGF)* | 42–5, 44, 51  
基础(bFGF, FGF-2) | 39, 43, 49–50  
*receptor 1 (FGF-R1)* | 43  
*type 1 (FGF-1)* | 42  
*type 5 (FGF-5)* | 43  
*fibroblasts in ACL autografts* | 82  
*tendon* | 56  
*heat-induced death* | 57, 64  
*in tissue repair* | 39–40, 64–5, 238  
*fibrocartilage* | 60, 63  
*fibrosis* | 26, 27, 238  
*five dot drill* | 293  
*flexibility* | 232–53  
*components* | 233–4  
*definition* | 233  
*effects of heat* | 207  
*injury risk and* | 156, 235–6  
*measurement* | 238–9  
*terminology* | 233–5  
*flexibility training (stretching)* | 232–3  
*definition* | 233  
*general benefits* | 235–7  
*decreased muscle soreness* | 236  
*decreased risk of injury* | 235–6  
*improved performance* | 236–7  
*methods* | 240–5  
*ballistic stretching* | 242–3  
*passive stretching* | 243–4  
*proprioceptive neuromuscular facilitation see proprioceptive neuromuscular facilitation static stretching see static stretching*  
*physiology* | 239–40  
*practical guidelines* | 253  
*proprioceptive effects* | 279  
*scientific basis* | 245–53  
*effect of stretch position* | 252–3  
*efficacy of session* | 248  
*efficacy of repeated sessions* | 248–9  
*frequency* | 246–7  
*optimum duration* | 245–6  
*warm-up prior to* | 249–52  
*sport-specific protocols*  
*running* | 295, 296  
*soccer* | 297, 298  
*tennis/baseball* | 292, 293, 294  
*fluidotherapy* | 209–10  
*fluori-methane* | 205  
*flurbiprofen* | 188  
*foot orthotics* | 304–7  
*formoterol* | 200  
*fractures* | 100–6  
*aim and principle of treatment* | 100–1  
*fixation methods* | 102–3, 104  
*healing* | 101–3  
*biological influences* | 103  
*mechanical influences* | 102–3  
*primary and secondary* | 101  
*management* | 103–6, 149, 212, 223–4  
*sports-related* | 100  
*Freeman (BAPS) board exercises* | 281, 284  
*frustration* | 165  
*functional biomechanical deficit complex* | 147  
*functional rehabilitation* | 288–97  
*galvanic stimulation* | 214  
*gap junctions* | 22, 23  
*gaseous cavitation* | 210, 211  
*gastrointestinal disorders* | 8  
*gastrointestinal effects*  
*NSAIDs* | 191  
*opioids* | 197–8  
*gate-control theory* | 218  
*gelatinase* | 16  
*gender differences, resistance training effects* | 263  
*gene therapy* | 50–1  
*glucagon* | 24–5  
*glucocorticoids see corticosteroids*  
*glucose metabolism* | 126–8  
*GLUT-4 transporter* | 126, 127, 136  
*glycogen, muscle* | 129  
*glycogen synthase* | 133  
*glycolytic enzymes* | 133  
*glycosaminoglycans (GAG)*  
*articular cartilage* | 110–8  
*ligament/tendon* | 65, 74, 76  
*goal setting* | 169, 172, 173  
*goals, rehabilitation* | 4, 145–6  
*acute phase* | 264, 288  
*functional phase* | 289  
*recovery phase* | 289  
*return to sport phase* | 269  
*training phase* | 265  
*golf* | 291  
*Golgi–Mazzoni corpuscles* | 279  
*Golgi tendon organ (GTO)* | 240, 241, 244  
*granulation tissue* | 26–7  
*grief reactions* | 164–5  
*growth factors*  
*bone* | 100  
*in fracture healing* | 101, 105  
*in muscle injury* | 39, 41, 42–5, 52  
*in osteochondral fractures* | 109  
*therapy* | 49–51, 91  
*growth hormone* | 24, 213  
*gymnastics* | 4, 5, 310  
*haemostasis* | 18  
*Hageman factor* | 21  
*hamstring flexibility* | 235  
*injury risk and* | 235  
*methods of increasing (stretching)* | 241–2, 244  
*stretching studies* | 246, 247, 248, 249, 252, 253  
*testing* | 238–9  
*hand*  
*injuries* | 311–12  
*orthoses* | 312  
*Haversian systems* | 99  
*healing see tissue repair/healing*  
*health services, international competitions* | 7–8  
*heart dimensions* | 122–4  
*heart rate*  
*detraining effects* | 120–1  
*maximal* | 121, 122  
*recovery* | 122  
*reduced training and* | 136  
*resting* | 120, 122  
*submaximal* | 120, 122  
*heat*  
*physiological effects* | 207, 234  
*transfer* | 206–7  
*heat therapy* | 206–13, 224  
*deep (diathermy)* | 207, 210–13  
*indications* | 207–8  
*prior to stretching* | 250  
*superficial modalities* | 208–10  
*heel pain* | 212, 215  
*helix–loop–helix family, basic (bHLH)* | 45, 47  
*heparin sulphate proteoglycans* | 51  
*hepatic enzyme growth factor (HGF)* | 44  
*hexagon drill* | 293  
*hexokinase* | 133  
*high-density lipoprotein cholesterol* | 127, 128  
*high-voltage pulsed galvanic stimulation (HVPGS)* | 214, 216  
*histamine* | 15, 18–19  
*Hoffman (H) reflex* | 152, 153  
*hormones* | 24–5  
*hot water bottles* | 208  
*H reflex* | 152, 153  
*Hubbard tanks* | 209  
*hyaluronate, intra-articular* | 74  
*hydrocodone* | 197, 198  
*hydrocollator packs* | 208  
*hydrocortisone phonophoresis* | 213  
*hydrocortisone* | 197  
*hydrotherapy* | 208–9  
*β-hydroxyacyl-coenzyme A (CoA) dehydrogenase* | 132–3  
*β-hydroxybutyrate dehydrogenase* | 133  
*5-hydroxytryptamine (5-HTP)* | 198  
*hyperalgesia* | 195
interferon-α, interferential current–220

intermediate model of psychological

insulin-like growth factor II (IGF-II)

insulin-like growth factor I (IGF-I)

insulin–24–5

inflammatory cells–15–20

inflammatory mediators–10–15

inflammation–10–27

infections, repeated pyogenic–22

indomethacin–188

immune system–15–22

immobilization–149–55, 258

cardiovascular changes 154–5

joint changes 73–4, 153–4

ligament tensile strength and

74–5

muscle changes 150–1

in muscle injury 49

neural changes 151–2

vascular changes 153

see also detRAINing; mobilization

immune system 15–22

indomethacin 188

injections, repeated pyogenic 22

inflammation 10–27

acute 22–5

chronic 10, 25

components 10–22

corticosteroid actions 192

in muscle injury 38–40, 49

neurogenic 195

NSAID actions 188–90

opioids and 195

process 22–5

in tendon/ligament injury 64–5

tissue repair 25–7

inflammatory cells 15–20

in muscle injury 38–41

in tissue injury 22

inflammatory mediators 10–15

injections, corticosteroid/local anaesthetic

insulin 24–5

action 126, 127, 136

insulin-like growth factor I (IGF-I) 39, 44

in fracture healing 101

therapeutic potential 49–50

insulin-like growth factor II (IGF-II) 39, 43, 44

in fracture healing 101

integrated model of psychological

responses 161, 162, 175, 176

interferential current 220

interferon-α (IFN-α) 12

interferon-γ (IFN-γ) 12, 25, 199

interleukin 1 (IL-1) 11, 12, 13

in fracture healing 101

in inflammation 25

in muscle injury 39, 41

interleukin 1α (IL-1α) 12, 24

interleukin 1β (IL-1β) 11, 12, 199

interleukin 2 (IL-2) 12

interleukin 3 (IL-3) 12

interleukin 4 (IL-4) 11, 12

interleukin 5 (IL-5) 12

interleukin 6 (IL-6) 11, 12, 199

in fracture healing 101

in inflammation 24, 25

in muscle injury 39, 41, 43–4

interleukin 8 (IL-8) 12

interleukin 10 (IL-10) 11, 12

internal rotation/external rotation ratio 298

international competitions 7–8

International Olympic Committee (IOC), prohibited substances 191–2, 195, 200

interoceptive sensation 274

intersegmental dynamic analysis 313

intervertebral septal wall thickness 122, 123

inverse stretch reflex 240, 244

iontophoresis 192, 215

isokinetic exercise 262, 266, 267

isokinetic testing 266, 267

isometric muscle action 261

joint

corticosteroid/local anaesthetic injections 193–4

effusion, reflex muscle inhibition 152, 153, 258

immobilization-induced changes 73–4, 153–4

sensory receptors 275, 276

joint capsule

effects of immobilization 73–4

mechanoreceptors 274

joint range of motion (ROM) 232–53

athletic performance and 236–7

decreased (contracture), after injury 153–4, 237–8

definition 233, 235

duration of stretching and 245, 246

effects of immobilization 153–5

end-point detection 239

excess (hypermobility) 232

frequency of stretching and 246–7

measurement 238–9

methods of increasing 240–5

retention after one stretching session 248

after repeated stretching sessions 248–9

termology 233–5

warm-up effects 249–52

see also flexibility

jumper’s knee 63, 310

Kabat technique see proprioceptive

neuromuscular facilitation

kallikrein 21

keratan sulphate 107

ketoprofen 188

ketorolac 188

kicking 291

kinaesthesia 274

kinetic chain 144–5, 280–1

exercises, closed and open 266, 281

muscular balance and 155–6

sport-specific 290–1

toward knee

anteromedial pain (jumper’s knee) 63, 310

immobilization, effects 74–5

injuries 6

limited motion casting 66, 78

mechanisms 59–60

proprioceptive effects 278–9, 310

strength deficits after 6, 258–9

orthoses (braces) 301, 302, 303, 307–10

sensory receptors 279, 310

see also anterior cruciate ligament;

medial collateral ligament, knee; posterior cruciate ligament

Krause bulbs 275, 276

lactate, blood 128–9, 136

lactate dehydrogenase (LDH) 48, 133

ladder drill 294

laser photostimulation 73

lateral epicondylitis (tennis elbow) 189, 311

acupuncture 199–200

counterforce orthosis 312

ultrasound therapy 212–13

left ventricular end-diastolic

dimension 122, 123

left ventricular mass 122, 123, 124

left ventricular wall thickness 122–3, 124

leptin 25

lethargy 25

leucocytes

in inflammatory response 22, 23

in muscle injury 42

in tissue repair 26

see also inflammatory cells;

macrophages; neutrophils
leukaemia inhibitory factor (LIF) 39, 43–4
levorphanol 197
lidocaine 193, 200
ligament 56–92
biomechanics 57–60
effects of immobilization 73, 74, 154
injury 148
- collagen fibril size distribution after 61–2
- early controlled mobilization 65–6, 78
grade 1 tears 59
grade 2 tears 59
grade 3 rupture 59
mechanisms 59–60
rehabilitation 148
insertion to bone 60
mechanoreceptors 274
repair/healing 64–9
- acute inflammatory phase 64–5
- effects of NSAIDs 189
- phases 64, 65
- proliferative phase 65–7
- remodelling and maturation phase 67–9
- tissue engineering approaches 91–2
structure 56–7, 58
tensile strength see tensile strength, ligament/tendon
ligamentous laxity 156
line drill 294
lipid metabolism 25, 126–8
lipoprotein(a) 127
lipoprotein lipase 127, 133
load–deformation (strain) curve
- Achilles tendon ‘fatigue failure’ 72
tendon/ligament 57–9
loading, mechanical
- ACL autografts 82
- articular cartilage injuries 108
- fracture healing and 102–3
- muscle regeneration and 51–2
- see also exercise; immobilization; mobilization
local anaesthetic injections 192, 193–4, 200
locus of control 168
loss, injury as 164
low back pain see back pain, low
low-density lipoprotein cholesterol 127
lower limb
- orthoses 304–10
- proprioceptive re-education 282–4
lymphocytes 19–20, 40
macrophages 15–16, 26
- ED1+ 41
- in muscle injury 40–1, 44
- phagocytosis 41
macrotia 146, 148
magnetic therapy 73, 223–4
major basic protein (MBP) 17–18
major histocompatibility complex (MHC) 19
malate dehydrogenase 132–3
march fracture 100
marrow stimulation procedures, cartilage injuries 110–11
mast cells 15
maximal oxygen uptake see V̇O₂max
mechanical factors, fracture healing 102–3
mechanoreceptors 274, 275, 276
meclomenamate 188
medial collateral ligament (MCL), knee
decorin antisense therapy 91–2, 93
meckel corpuscles 275, 276
membrane attack complex (MAC) 20–1
meniscectomy 196
meniscal resection, referral to 174
meperidine 196, 197, 200
mepivacaine 200
meprobamate 197
Merkel’s disks 240, 275, 276
mesenchyme stem cells 91
metabolic adaptations
- effect of detraining 125–9, 137–8
- reduced training and 136
- metabolic rate, resting 128
- metatarsal fracture 100
- met-enkephalin 195
methadone 197, 200
microfracture technique, cartilage injuries 110–11
mitochondria
- ATP production 134
muscle
- activation, impaired 258–9
- agonist–antagonist ratios 260, 298
- atrophy/hypotrophy/wasting 259
- assessment 259
- electrical stimulation 215–17
- immobilization-induced 151
- capillarization 130, 153, 260
- circumference measurement 259
- effects of detraining 129–34, 138
- electrical stimulation 214, 215–18, 262, 265
- endurance 260, 269
- enzymatic activities 132–3, 136, 260
- force development, speed of 261
- glycogen 129
- healing/repair 37–52, 149
- biological treatment approaches 48–52
- degeneration phase 25, 37, 38–41
- maturation phase 37, 45–8
- regeneration phase 35, 37, 41–5
imbalance, relative 155–6
immobilization-related changes 150–1
injury 35–52, 148–9
- biomechanics 62–3
- in situ 35
- shearing 35
- mass
- detraining effects 131–2
- immobilization-induced loss 151
- reduced training and 136
- myoglobin 132
- power training 268–9
- reduced training and 136
- re-education 216
- reflex inhibition 152, 216, 258–9
- sensory receptors 275, 276
soreness, delayed-onset see delayed-onset muscle soreness
specific proteins 45–8
strength see strength, muscle structure 35–7
volume
postinjury reduction 259
resistance training effects 261, 265
muscle fibre (myofibre)
damage 35, 38
distribution, detraining effects 130–1
fast-twitch (FT, type II)
detraining effects 130–1, 132
effects of injury 259
myosin isoforms 46
myotendinous junction 63
regeneration 47
maturity 45–8
regeneration 35, 37, 41–5
size
detraining effects 131–2
immobilization-induced losses 151
slow-twitch (ST, type I)
detraining effects 131, 132
effects of injury 259
myosin isoforms 46
regeneration 47
muscle relaxants 194
muscle–tendon–bone injury 62–3
muscle–tendon junction injury 148–9
biomechanics 63, 64
tissue repair 67
musculotendinous unit
elastic energy storage 57, 236–7
flexibility see flexibility
length 233, 235, 238
ROM see joint range of motion
stiffness 234
tension 233–4
myeloperoxidase 16
Myf5 45
myoblasts
differentiation into 42, 44–5
fusion 45
myocardial ischaemia, silent 124
MyoD 43, 45, 47, 48
myofascial trigger points 205, 212, 218
myofibre see muscle fibre
myofibroblasts 238
myogenic precursor cells 36, 41–2
44–5
myogenic regulatory factors (MRFs) 36, 45, 47
myogenin 45, 47
myoglobin, muscle 132
myosin 46–7
myosin heavy chain (MHC) isoforms 46–7
myosin light chains (MLC) 46, 47
myostatin 39
myotendinous junction injury see muscle–tendon junction injury
myotubes 45
nabumetone 188
naproxen 188
narcotics see opioids
natural killer (NK) cells 20
nerve endings, free 276, 279
nerve growth factor (NGF) 39, 43, 44
nerve injuries, upper extremity 312
nerves
afferent 275
electrical stimulation 214, 218–20, 262
neural activation
electrical stimulation and 262
flexibility training 239–40
immobilization and 151–2
neuroendocrine system 24–5
neurogenic inflammation 195
neuromuscular activation 260–1
assessment 260–1
effects of injury 260
neuromuscular re-education 274
neuropaxia 204
neutropenia, congenital
dysgranulopoietic 17
neutrophils 14, 16–17
in inflammatory response 22
in muscle injury 39–40
in tissue repair 26
nociceptors, peripheral 195, 196
non-steroidal anti-inflammatory drugs (NSAIDs) 187–91
dosing 188, 190–1
effects on healing 41, 49, 189, 190
history 187–8
mechanism of action 13–14, 15, 23, 188–90
side effects 191
topical 190–1
non-union, fracture 103
NSAIDs see non-steroidal anti-inflammatory drugs
oedema
electrical stimulation and 218
pathophysiology 14, 23
omeprazole 191
opiate pain control theory 218
opioid receptors 195
opioids (narcotics) 194–8, 200
clinical applications 196
dosage 197
endogenous 195, 198
side effects 196–8
orthoses 301–13
accommodative 302
ankle 302, 307
biomechanics 302–3
clinical application 303–4
corrective 302
custom made 303, 304–6
dynamic 302
fabrication 303
foot 304–7
functional 302
future research 312–13
knee 301, 302, 303, 307–10
low back 310–11
prophylactic 302
rehabilitative 302
static 302
terminology and classification 301–2
upper extremity 311–12
osteoarthritis 223
osteoblasts 99–100, 101, 102
osteochondral autograft transplantation 111
osteochondral fractures 108, 109
osteoclasts 99–100, 101
osteocytes 99
osteoprogenitor cells 100
overload syndromes, chronic 148
overuse injuries 4–5, 148
see also chronic injuries
oxaprozin 188
oxidative enzymes, muscle 132–3, 136
oxycodone 197, 198
oxygen difference, arteriovenous 132
oxygen uptake, maximal see $V_{O2max}$
Pacinian corpuscles 240, 275, 276, 279
pain management 204–24
see also analgesia
mechanisms 14
pathways 199
paracetamol (acetaminophen) 198
paraffin baths 209
parasympathetic tone 124
parathyroid hormone 105
passive stretching 243–4
patellar dislocation/subluxation 310
patellar taping 310
patellar tendinopathy 63, 310
patellar tendon
ACL autografts
goat studies 85–90
human studies 79–81, 82
biomechanics of injury 63
patellar tendon (cont.)
collagen fibril size distribution
  goat 85, 86
  human 78, 79
  rabbit 61, 62
  healing 68
  insertion 60
  origin 60
patellofemoral pain syndrome 259, 310
pathophysiology of injuries 10–27
PCL see posterior cruciate ligament
pentazocine 197
percutaneous electrical nerve stimulation (PENS) 220
performance, athletic
  flexibility and 236–7
  training cessation and 134–5
peripheral resistance, total 122, 123, 124
peroxidase 16, 17
personal factors
  psychological responses to injury 166, 168
  rehabilitation outcomes 170
  personality 170
phagocytosis, in muscle injury 41
pharmacological agents 187–98
phosphofructokinase 133
phosphorylase 133
physical modalities
  phonophoresis 192, 213
  warm-up before 250, 251
  upper limb 281–2
  lower limb 282–4
  proximal interphalangeal joint, dislocation 311, 312
plasmin 21
plasma volume 119–20, 122
plasmin 21
plasticity
  Achilles tendon 72
  effects of stretching 240
plate fixation, fractures 103
platelet–activating factor (PAF) 14, 22
platelet–derived growth factor (PDGF) 39–40, 43, 101
platelets 18–19, 22, 27
plyometrics 268–9
plyometrics/agility protocols
  running 295, 296
  soccer 297, 298
  tennis/baseball 292–3, 294, 295
PNF see proprioceptive neuromuscular facilitation
postcapillary venules 23
posterior cruciate ligament (PCL)
  effect of mobilization 75, 76
  mechanisms of injury 60
prednisolone iontophoresis 214–15
prednisone 192, 193
prehabilitation 289
  exercises 291–2
  protocols for specific sports 292–8
preparticipation exam 7
prevention of injury model 3–4
preventive rehabilitation 7
procaine 200
propoxyphene 197, 200
proprioception 274–86
  definition 274
  effect of injury 278–9, 310
  measuring performance 277–8
  neurological mechanisms 274–7
  re-education 279–84, 285–6
  adaptation to training 284
  lower limbs 282–4
  upper limbs 281–2
proprioceptive neuromuscular facilitation (PNF, Kabat method) 240–2, 280–2
  contract–relax (CR) 241
  contract–relax–agonist contract (CRAC) 241–2
  efficacy of one session 248
  efficacy of repeated sessions 249
  number of repetitions 247
  optimum duration 246
  warm-up before 250, 251
proprioceptive reflexes 276
proprioceptive sensation 274
proprioceptors 275, 276
prostaglandin E2 (PGE2) 13, 14
proximal interphalangeal joint, dislocation 311, 312
psychological factors 160–77
  future research directions 175–6
  implications 173–5
  in injury occurrence 176
  in rehabilitation outcomes 169–73
psychological interventions 172–3, 175–6
psychological responses to injury 161–9
  behavioural 167–9
  emotional 164–7
  explanatory models 161–4, 175, 176
psychologists, sports 174–5
psychotherapy 174
pyramidal system 277
quadriceps–hamstrings ratio 260, 298
quadriceps muscle
  reflex inhibition 258–9
  strength differences between limbs 259
  wasting after injury 259
quadriceps tendon, insertion 60
racquet stretch 293
range of motion (ROM) see joint range of motion
rectus femoris tendon injury 64
reflexes, proprioceptive 276
reflex inhibition 152, 216, 258–9
rehabilitation 145–6, 264–70
  adherence 167–9, 172
  first (acute) phase 264–5, 288
  functional 288–97
  functional phase 288, 289
  goals see goals, rehabilitation
  lack of effective 4, 6
  phases 288
  plan 45
  preventive 7
psychological influences on outcome 169–73
recovery phase 288, 289
second (training) phase 265–9
  third (return to sport) phase 269–70
reinjury risks 145, 270
relaxation 172, 173
relaxation syndrome 123
renal toxicity, NSAIDs 191
repetition maximum (RM) 264, 265
  repetitions, number of 264, 265, 269
resistance/strength training 260–3
  children and adolescents 263
  contralateral effects 261
  dynamic 261–2
  electrical muscle stimulation and 262
endurance training interaction 264
  gender differences 263
  methods 261–2
  muscle volume increase 261
  neuromuscular activation 260–1
  prehabilitation protocols
  running 295, 296
  soccer 297, 298
  tennis/baseball 292, 294, 295
in rehabilitation 264, 265–8
  static 261, 262
respiratory adaptations
  effects of detraining 117–25
  reduced training and 136
respiratory depression, opioid-induced 197
respiratory disorders 8
respiratory exchange ratio (RER) 125, 136
respiratory tract infections, upper 20
rest 149–55
see also detraining; immobilization
rest, ice, compression and/or elevation (RICE)
in ligament/tendon injuries 66, 67
in muscle injuries 49
return to sports 269–70, 289
criteria 298–9
functional programmes 299, 300
RICE see rest, ice, compression and/or elevation
rofecoxib 190
role, identification with athlete 166, 170
ROM see joint range of motion
rotator cuff disorders
acupuncture 199
diagnosis 146–7
physical modalities 212, 215
Ruffini corpuscles 275, 276, 279
runner’s knee 310
running 291
prehabilitation 295–6
return to sports 298
salbutamol 200
salicylic acid 187–8
salmeterol 200
satellite cells 35–7, 52
activation 37, 42–4
ageing changes 37
differentiation into myoblasts 42, 44–5
growth factors/cytokines affecting 39
myogenic regulatory factors 36, 45
scaphoid fractures 311, 312
scar tissue 238
decorin antisense therapy and 91–2, 93
effect of stretching 240
in ligament/tendon healing 66, 67
sedation, opioid-induced 197
self-talk 172, 173
sensation, classification 274–5
sensory receptor organs 275, 276
serotonin 15, 18–19
severity of injury 7
shoe, athletic 306
shoulder
injuries 6, 146–7, 212
prehabilitation 292–5
situational factors
psychological responses to injury 166–7, 168, 169
rehabilitation outcomes 170–1
skeletal dysfunction, flexibility and 237
soccer 291
prehabilitation 296–8
return to sport programme 299, 300
social support 170–1
sodium salicylate 188
somatosensory cortex 276
somnolence 25
spondylolysis 149, 311
sport profiling 289–90, 291, 299
sports injuries 4–7
acute see acute injury
anatomical distribution 6–7
benefits 176
chronic see chronic injuries
diagnostic context 146–8
epidemiology 3–8, 145
flexibility in prevention 235–6
frequent diagnoses 7
multifactorial nature 146
pathophysiology 10–27
physiological/functional effects 144–57
psychological responses 161–4
reinjury risk 145, 270
severity 7
time from onset 5–6
types 4–5
sports psychology professionals 174–5
sports-specific demands 289–92
critical point framework 288
prehabilitation exercises 291–2
sprains 7, 8
spray and stretch technique 205
sprinting 291
static stretching 243–5
efficacy of one session 248
efficacy of repeated sessions 248–9
number of repetitions 246–7
optimum duration 245–6
position effects 252–3
practical guidelines 253
repeated 244–5
by therapist (passive) 243–4
warm-up prior to 249–52
active warm-up 250–2
passive warm-up 250
stem cells, mesenchyme 91
steroidal medications 191–4
see also anabolic steroids; corticosteroids
stiffness
after ACL injury 69
definitions 233, 234
straight leg raise (SLR) 239
strain, mechanical, fracture healing and 102
strains 7, 8
strength, muscle 258–70
cross-transfer effect 137, 261
detraining effects 135, 138
difference between two limbs 259
electrical muscle stimulation and 217–18
immobilization-related loss 150–1
postinjury reduction 6, 258–60, 270
reduced training and 136
training see resistance/strength training
stress fractures 100
neuroendocrine response 24–5
psychological, injury-related 161
stress inoculation training 172, 173
stress relaxation
definition 233, 234
procedure 245, 246–7
stretching see flexibility training
stretch reflex 239–40, 242, 244, 276
inverse 240, 244
stroke volume 121, 122
ST segment depression 124
subclinical adaptation complex 147
substrate availability/utilization 125–8
succinate dehydrogenase 132–3
suicide risk 165
sulindac 188
superficial digital flexor tendon, equine 64
supraspinatus tendon 64, 68
swimming 7, 291
sympathoadrenergic tone 120, 124
synovial fluid, immobilized joints 154
synovial hyperplasia, immobilized joints 154
systemic lupus erythematosus (SLE) 22
temperature
infrasynovial, after warm-up 251
intratendinous 57, 64
tissue, flexibility and 249–50
tendon 56–92
biomechanics 57–60
corticosteroid injection 194
hyperthermia 57, 64
injury 146, 148
biomechanics 64
collagen fibril size distribution after 61–2
rehabilitation 148
ultrasound therapy 212
insertion to bone 60, 63
tendon (cont.)
mechanoreceptors 274
repair/healing 64–9
acute inflammatory phase 64–5
phases 64, 65
proliferative phase 65–7
remodelling and maturation phase 67–9
tissue engineering approaches 91–2
rupture, spontaneous 64
structure 56–7, 58
tensile strength see tensile strength, ligament/tendon
tendonitis (tendinosis) 7, 8
tennis 291
prehabilitation 292–5
upper limb injuries 311, 312
tennis elbow see lateral epicondylitis
tensile strength ligament/tendon 74–8
ACL auto/allografts 82, 84, 85
collagen fibril size distribution and 61, 62, 82, 84
effect of immobilization 74–5
effect of mobilization 65–6, 75–8
postsurgery 69
surgical repaired ligaments 76–8
muscle, recovery after injury 149
tension, musculotendinous unit 233–4
TENS (transcutaneous electrical nerve stimulation) 218–20
terbutaline 200
testosterone 131, 136, 192
testosterone:cortisol ratio 131, 136
TGF-β see transforming growth factor beta
thermotherapy see heat therapy
thromboxane A₂ (TxA₂) 13, 18–19
throwing 291
return to sports programme 298, 299
thyroid hormone 47
thyroid-stimulating hormone 24
tibialis anterior tendon, ACL allografts 82–3
tibialis posterior tendon, ACL allografts 83
tibial plateau fracture 100, 103, 104
tissue engineering, ligament/tendon repair 91–2
tissue injury
biomechanics 62–4
complex 147, 148–9
inflammatory response 22–5
tissue overload complex 147
tissue repair/healing 25–7
bone 100–6
cartilage 108–12
effects of NSAIDs 41, 49, 189, 190
electrical stimulation and 214
physical modalities and 73
skeletal muscle 35–52
soft tissue contracture after 237–8
tendons and ligaments 64–92
T-lymphocytes 14, 19, 20, 26
tolerance, opioid 198
tolmetin 188
total peripheral resistance 122, 123, 124
trabecular bone 99
traction 220–3, 224
cervical 221–2
lumbar 222–3
training
adaptations, retention 135–7
cessation see detraining
coordination 284–5
cross- see cross-training
endurance, postinjury 263–4
reduced 136–7, 138, 155
unilateral, cross-transfer 137, 138, 261
see also flexibility training; resistance/strength training
transcutaneous electrical nerve stimulation (TENS) 218–20
transferrin 24
transforming growth factor beta (TGF-β) 11, 12
in fracture healing 101, 102
in muscle injury 39, 43, 51, 52
traumatic injuries 4–5
triglycerides 127
triacontanol acetate 193
triiceps surae stretching 250
trigger points, myofascial 205, 212, 218
triglycerides 127
tumour necrosis factor alpha (TNF-α) 11, 12, 13, 14
in inflammation 24, 25
in muscle injury 39, 41
in tissue repair 26
tumour necrosis factor beta (TNF-β) 12
ulnar neuropathy 312
ultrasound therapy 210–13
application methods 210–11
clinical applications 211–13
fracture healing and 103–5
prior to stretching 250
soft tissue repair and 73
see also phonophoresis
unstable planes, exercises on 280, 281, 284–5
upper limb
orthoses 311–12
proprioceptive re-education 281–2
vapocoolant sprays 205
vascular cell adhesion molecule 1 (VCAM-1) 42
vascular permeability 23
vasculature
immobilization-related changes 153
injury 18–19
response to injury 22–3
vasoconstriction 23, 27
vasoconstrictor agents 200
vasodilatation 23
vasopressin 24, 25
ventilatory function 124–5
very late antigen 4 (VLA-4) 42
very low-density lipoprotein cholesterol 127
viscoelasticity 233–4, 235
ACL autografts 88–9
viscous (properties) 233, 234
arteriovenous oxygen difference 132
and blood volume and 119–20
effects of detraining 118–19, 154–5
reduced training and 136
stroke volume and 121
volleyball 291, 312
von Willebrand factor 18
warm-up
active 250–2
passive 250, 252
prior to static stretching 249–52, 253
tendon/ligament changes 59
warm water immersion 209
water
in articular cartilage 107–8
exercise in 208–9
in ligament/tendon repair 65, 67
warm, immersion 209
weightbearing in chondral injuries 111
early, in fractures 102–3
weight-training 265
circuit 270
closed and open kinetic chain 266
free weights vs weight machines 266–8
see also resistance/strength training
whirlpools 209
wound contraction 238
wrist
injuries 311, 312
orthoses 312