CHAPTER SAMPLER

Routledge Free Chapter Collection for BASES Members

Routledge
Taylor & Francis Group

www.routledge.com

Contents

CHAPTER 2

QUALITATIVE BIOMECHANICAL ANALYSIS OF TECHNIQUE

Adrian Lees

INTRODUCTION

Qualitative analysis is a method used to evaluate technique in the performance of sports (or exercise) skills. It uses observation and can be supplemented with a visual recording, such as video, with playback facilities such as single frame, slow motion and repeated viewing. It relies on a knowledge of the relevant sport and sports skill, as well as a knowledge of 'principles of movement'. Finally, qualitative analysis should be conducted using a chosen analysis model.

It is important to distinguish between technique and performance. Technique is defined in general terms as the 'way of doing' and the technique used in the performance of a specific sports skill can be defined as the way in which that sports skill is performed. Performance is defined in terms of outcome, and technique is only one of several factors that influence outcome. This distinction has been difficult for biomechanists to make due to the popularity in recent years of quantitative analysis and attempts by investigators to undertake a 'biomechanical analysis of performance' which, in many cases, attempts to measure aspects of technique. A simple example can illustrate the difference. In a sprint, an individual may be able to demonstrate the technique (way of doing) of an elite sprinter but not be able to achieve the same level of performance. This is because the individual lacks the physical, physiological, psychological or other attributes required for a higher-level performance. Thus, good performance requires good technique, but good technique does not guarantee good performance.

Good technique is a prerequisite of good performance so it is sensible to attempt to understand what good technique means for individual sports skills. The advantage of the qualitative analysis approach is that an analysis can be undertaken quickly, without recourse to expensive, time consuming and often restrictive methods characterised by contemporary quantitative methods. A further advantage is that a sports skill can be analysed holistically; something that is quite difficult to do using quantitative methods.

Qualitative analysis is not a new approach to the analysis of sports skills. In fact this was the main approach in the 1950s and 1960s, mainly due to increasing interest in the techniques used to perform sports skills but the lack of equipment and methods to investigate them (see Lees, 2002, for a review). Out of these early studies evolved what we now recognise as qualitative analysis. Unfortunately, these approaches were overtaken by the thrust in biomechanics for quantitative analysis; a suitable framework for qualitative analysis was never developed. Researchers have subsequently defined their own frameworks but no generally agreed approach has emerged. In this chapter I provide the basis of qualitative analysis, starting with analysis models, continuing with principles of movement, and concluding with some contemporary thoughts and developments on this topic.

ANALYSIS MODELS

The phase analysis model

Phase analysis is the descriptive process of dividing up a movement into relevant parts so that attention can be focused on the technique used in each part. Some authors identify three main phases to a skill (*retraction, action and follow-through*), while others identify four main phases (those above but preceded by a *preparation phase*), a recognition that the start of a movement may also influence the way a movement is performed (e.g. in a soccer kick the distance away from the ball and angle of approach are relevant factors to performance; in golf, the stance, how the club is held and position of the feet relative to the ball are also relevant factors). Some authors identify more than four phases but these are often sub-phases of the four main phases mentioned earlier. Most authors acknowledge that the phases can be further broken down into *sub-phases* and that the distinction between one phase or sub-phase and another is arbitrary and determined by the particular skill and the needs of the analyst. Nevertheless, this process of breaking a skill down into its functional parts is an important first analytical step.

The *preparation phase* describes the way in which the performer sets or prepares for the performance of the skill. For example, as noted earlier, it may relate to the start position and/or the way the ball is placed in a soccer penalty kick, or the way the club is held and/or ball placed on the tee in a golf drive. The *retraction phase* refers to the withdrawal of, typically, the arm or leg prior to beginning the main effort of performance. For example in kicking, the kicking leg is drawn back; in tennis the retraction phase is represented by the backswing. The *action phase* is where the main effort of the movement takes place. For example, forward motion in executing a tennis serve, or the throwing action when propelling a javelin. The *follow-through phase* allows the movement to be slowed down under control and is thought to be necessary to avoid injury that might occur with rapid limb deceleration. Some examples are given in Table 2.1. As noted earlier, the phase analysis model requires the phases and sub-phases, if appropriate, to be identified and described. Some examples of sub-phases are given in Table 2.2.

Table 2.1 Examples of the four phases from selected skills.

Event/skill	*Preparation*	*Retraction*	*Action*	*Follow-through*
Long jump	Run up	Body adjustments during penultimate stride to touchdown	Take-off, from touchdown to take-off	Flight and landing
Golf swing	Correct grip and stance	Backswing	Downswing	Follow-through
Penalty kick	Run up	Retraction of kicking leg	Forward swing of kicking leg and contact	Follow-through

Table 2.2 Examples of the sub-phases for the action phase of selected skills.

Event/skill	*Action phase*	*Sub-phase 1*	*Sub-phase 2*	*Sub-phase 3*
Long jump	Take-off	Compression (knee flexion)	Extension (knee extension)	-
Golf swing	Downswing	Weight shift initiation of movement with the hips	Wrist locked arms rotate as single unit	Wrist unlocks and arms rotate as double unit
Penalty kick	Forward swing of kicking leg	Hips rotate forwards and leg flexes at knee	Knee extends to contact	-

The description of phases and sub-phases should identify *key moments* and *critical features*. Key moments are those points in time at which an important action is performed related to the 'way of doing'. One important key moment in striking sports is impact. However, events such as toe-off, foot-strike, maximum knee flexion, or minimum elbow angle, to name a few, would all be actions that define a key moment of the technique. Critical features are observable aspects of a movement, and refer to body or limb position (e.g. when catching a ball – crouch with arms and legs flexed; for a tennis serve – the 'back-scratch' position of the racket in the backswing) and motions (e.g. when catching a ball – give or retract the hands with the ball). It is worth noting that critical features are often related to coaching points and often are expressions of selected underlying movement principles (dealt with later). To complete the phase analysis model, the movement principles associated with the phase or sub-phase, key moments and/or critical features need to be identified. Movement principles are dealt with separately later in the chapter. The phase analysis model is given schematically in Table 2.3.

The performance outcome model

An alternative approach to the phase analysis model involves an analysis of the factors that influence performance. By focusing on those factors, it is claimed

Table 2.3 Schematic template for a phase analysis model.

Phase description	*Preparation*	*Retraction*	*Action*	*Follow-through*
Sub-phase description[1]				
Key moments[2]				
Critical features[3]				
Movement principles[4]				

[1] Each box under each phase should contain a brief description of the phase or (if appropriate) the sub-phase.
[2] Key moments are often related to the start and end of a phase or sub-phase.
[3] Critical features are often related to key coaching points.
[4] Movement principles can be identified by their abbreviated title (see text).

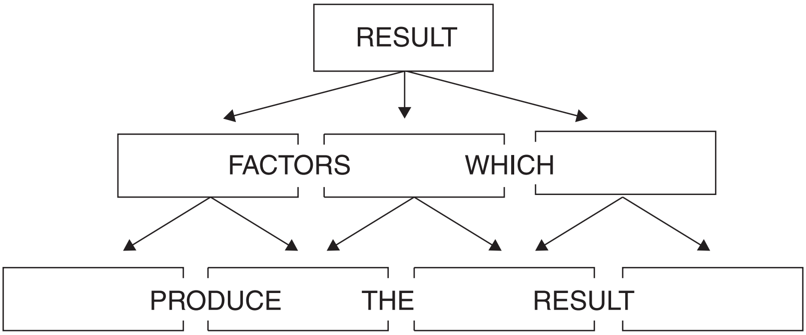


Figure 2.1 The Hay and Reid performance outcome model.

that faults and limiting factors in performance can be identified. The most influential of these models was proposed by Hay and Reid (Hay and Reid, 1982). The model was first developed for use in qualitative analysis but has also found widespread use in quantitative analysis to assist biomechanists in identifying important variables for quantification (Chow and Knudson, 2011). The model does not address aspects of technique (the way of doing) directly, but the mechanical relationships that govern performance. In that sense, the model is more closely linked with movement principles and one may view the model as a more direct and systematic approach to the identification of the mechanical principles that govern performance. The model is constructed as a hierarchy of factors on which the result (outcome) of the performance is dependent (Figure 2.1). The rule that governs the construction of a model for a particular skill is that each of the factors in the model should be completely determined by those factors that appear immediately below it either by (1) addition or (2) known mechanical relationships.

As noted earlier, the performance outcome model does not address issues of technique directly. For example, in a golf drive, the model will tell us that the speed of the club head must be high at impact, but not how to achieve it. Information on how to use the arms and club as a two-lever system, weight shift and hip-shoulder rotation are beyond the scope of the model. In other words, the model is able to identify factors relevant to performance but not aspects of technique relevant to these factors. Nevertheless, the Hay and Reid model would appear to be valuable for identifying a range of factors influencing performance and providing a framework from which technique can be discussed. In this sense, it is not an alternative model for technique analysis but is complementary to other methods.

MOVEMENT PRINCIPLES

The earliest and perhaps the most widely used scientific approach to the evaluation of technique is the application of *mechanical* principles. These have been articulated and developed over time and have tended to be referred to as 'biomechanical principles of movement' or *movement principles* for short. In general, these are a combination of principles based on mechanical relationships, multi-segment interactions and biological characteristics of the human musculoskeletal system. A movement principle is a description of how to achieve a specific movement outcome based on sound mechanical and/or biological principles. A movement principle is applied, in general terms, to help to understand how sports skills are performed.

Movement principles can be classified according to the general outcomes that they are associated with. These outcomes are speed production, force production, movement coordination and some that relate to special circumstances. Speed and force principles are based on mechanical relationships, such as those described by Newton's second law, and principles, such as the conservation of momentum. The equations underpinning these principles are not detailed here, to preserve the 'qualitative' nature of this chapter, but can be found in any good biomechanics text. The coordination principles are based on multi-segment interactions and have mechanical or biological foundations. This group of principles reflect the more complex operation of the human body. The inclusion of biological principles, such as the stretch-shorten cycle, is an acknowledgement of the biological factors that determine complex human movement. This is an area where a further development of movement principles would be warranted. The principles applied to special circumstances are where a number of interacting mechanical factors are commonly encountered, such as in the flight of projectiles, or where other phenomenon are frequently observed, for example the speed-accuracy trade off, reflecting a movement control limitation. There is no general agreement as to the number, or even the names, of movement principles. It is possible to identify principles that are related to specific performances. Some authors have identified as many as 53 movement principles, while others as few as six (Lees, 2002). The higher number tends to reflect the specific mechanical principles, whereas the lower number tends to reflect the more general principles (e.g. Bartlett, 2007). There have been some attempts to reduce the larger number of principles to a manageable form as 'core concepts' (Knudson, 2002); these contain a mixture of mechanical, multi-segment and biological principles

Table 2.4 Table of movement principles with codes.

Speed	*Force*	*Coordination*	*Specific Performance*
S1	F1	C1 AR	P1
S2	F2 ROM	C2 PDS	P2
S3	F3	C3	
S4 EPS	F4	C4 SSC	
	F5		
	F6		
	F7		

and provide some use in a practical context. Before illustrating how these are applied, it is necessary to identify what they are and give examples of how each can be used. Following is a list of those principles relevant to most sports and suitable for a qualitative biomechanical analysis of performance; they are outlined in Table 2.4. Although not definitive, they are thought to cover most applications.

Speed (S) principles

Speed principles are a range of movement principles which relate to the generation of speed in a part of or the whole of the body.

(S1) Whole body running speed

Running speed is increased gradually through sequential drives of the legs. The increase in speed, with each stride, is greatest at low speeds and reduces as whole body speed increases. Maximum running speed is achieved after about 40–50 m of sprinting. Therefore, to reach maximum speed, a person must be able to sprint for at least this distance. Some consequences of this are: (1) running speed is often controlled by the performer, due to the complexity of the skill and/or the high forces involved (in these cases the speed is kept sub-maximal in order to complete the skill), and (2) in field games, players are unlikely to reach their top running speed over a 10–15 m sprint, therefore their ability to accelerate is important.

(S2) Whole body rotational speed

Rotational movements of the whole body are completed more rapidly by bringing the limbs closer to the body's axis of rotation. For example, in a somersault the trampolinist rotates more rapidly when tucked; in a pirouette, the ice skater rotates more rapidly when the arms are brought close to the body; in a squash backhand stroke, the action phase begins with the racket arm close to the body so that the whole body can rotate into the movement before extending the racket arm. Conversely, extending the limbs slows the body down. This is done not only to slow down the rotation but to allow the performer more time to make a good landing. For example, stretching the arms above the head after a gymnastic vault slows the forward rotation of the gymnast; opening out after a tucked somersault slows the rotation of a trampolinist.

(S3) Limb rotational speed

To rotate a limb (e.g. arms or leg) rapidly requires the limb to be flexed and held close to the body. For example, in sprinting the leg is flexed tightly during the recovery part of the cycle. In a tennis serve the racket arm is flexed to achieve a back scratch position (a critical feature); in golf the arm and club are in a flexed position at the top of the downswing.

(S4) End point speed (EPS)

A high end point speed requires a large distance from the axis of rotation to the end point. Consequently at impact or release, the limb is at or close to full extension. For example, when striking in sports like tennis, golf, baseball and cricket, the action phase begins with the limb and implement held close to the body and as the phase develops, the end point (hand, foot, racket head, club head) is allowed to move away from the axis of rotation, thus increasing its radius of rotation.

Force (F) principles

Force principles are a range of movement principles which relate to the generation of force used to achieve a specific movement outcome.

(F1) Maximum force production

To produce the maximum effective force, a firm base is required against which to push. For example, in throwing events such as shot and javelin, the action phase occurs when the delivery foot is in firm contact with the ground; in jumping events such as long, triple and high jump, a firm surface is always used to push from. Conversely, if the surface moves, the effective force produced is reduced. For example, soft surfaces, such as turf and sand, are more difficult to run and jump on due to the deformation of the surface. In tennis serving, the server is often seen to come off the ground. Is this an exception? Not really. To gain maximum racket head speed the server extends the legs (generating maximal effective force) and it is only after this that the player throws the racket head towards the ball. This combination of vigorous movements directed vertically causes the body to lift off the ground. Other principles are also used in the performance of this skill in order to achieve high end point speed (see C2 and C4 sections that follow).

(F2) Range of motion (ROM)

The greater the limb's range of motion, the longer a muscle force can be applied for. Consequently, there is a possibility of achieving a greater effect by contracting muscle over a greater range of motion. For example, in running, as speed increases, the stride length increases, which occurs due to the greater extension (i.e. range of motion) of the leg during the drive-off. One implication of this principle is that if the joint has greater flexibility, it is likely that this will allow a greater force producing range of motion, leading to enhanced performance. Consequently there is thought to be a performance aspect to flexibility training.

(F3) Change of running direction

A change in direction of motion when running is produced by applying a force at right angles (perpendicular) to the current direction of motion. For example, in a side step or swerve made by a player in field games, the foot is placed so as to maximize the friction force applied to the surface. This friction force should be directed perpendicular to the current direction of motion and *not* towards the intended direction of motion.

(F4) Impact – stationary ball or object

When hitting a stationary ball or object, the implement making the impact must move in the direction it is intended that the ball or object being hit should go. For example, when taking a penalty kick, a goalkeeper can sometimes guess correctly the direction of the ball by carefully watching the motion of the kicker's foot as it moves to strike the ball.

(F5) Impact – moving ball or object

When hitting a moving ball or object, the striking implement must move in such a direction so as to take into account the motion of the moving object. It will always be the case that the direction of the striking implement at impact will be different from that which the ball or object subsequently goes. This divergence is related to the mass and speed of the two objects respectively (and another expression of principle F3). If, when heading a moving ball or making a cricket, baseball or tennis shot, the implement is swung to drive it in the direction of intended motion, the ball will *not* travel in this direction.

(F6) Stability

Objects are more stable if they have a wide base and low centre of mass. For example, in wrestling, at the start of a competition wrestlers spread their legs and lower their centre of mass to provide a stable base. Stability is a 'statics' concept. Beware of 'dynamic stability' which is a 'dynamics' concept beyond the scope of this chapter and probably not appropriate for a qualitative approach.

(F7) Resistance to motion in fluids

Resistance to motion when moving through air or water is reduced by decreasing the area of the body or object presented to the on-coming air or water (known as the cross-sectional area) and making a more streamlined shape. For example, in cycling, handlebars which are lowered and extended forward enable the cyclist to adopt a smaller cross-sectional area and more streamlined shape. Conversely, resistance is increased by increasing the cross-sectional area and making the shape less streamlined. For example in swimming, the area of the hand, and therefore the resistance produced by the arm pull, can be increased by using a hand paddle.

Coordination (C) principles

Coordination principles are a range of movement principles which relate to the coordination of motion between segments so as to achieve maximal or optimal performance.

(C1) Action–reaction: simultaneous movements of opposing limbs (AR)

The movement of one limb or body part helps the movement of the opposite (or contralateral) limb or body part. For example, in walking, running and sprinting, as one leg comes forward the contralateral arm also comes forward; an effective sprint start is one in which the arms drive vigorously to aid the force generation of the opposite leg. A good coaching point is that 'the arms drive the legs'; in hurdling (crossing over the hurdle) the lift of the lead leg is helped by bringing the opposite arm forward as far as possible; in pike movements in gymnastics, trampoline and diving, the performance of the pike is aided if the upper and lower body are brought together at the same time; the movement of one is helped by the movement of the other; when heading a soccer ball the same pike movement is produced which increases the speed with which the head makes contact with the ball.

(C2) Proximal-to-distal sequence of movements (PDS)

This is used when producing high-speed movements. Many skills require a coordinated sequence of rotational movements to achieve a high end point velocity. This is achieved by rotating the large segments close (proximal) to the body first and terminating in the rotation of the segment farthest (distal) from the body. For example, in most throwing/striking/kicking skills the action often starts with a step forward with the leg contralateral to the throwing/striking/kicking arm or leg, which has the effect of opening the hips. It continues with the hips rotating forward and is then followed by the trunk, shoulders and finally the elbow, hand and implement. The rotation speed of the earlier segment is built upon by the next segment, so as to build-up the speed of the end point sequentially. Rotation of the distal segments cause the end point to move away from the axis of rotation, thus increasing the distance of rotation (see speed principle S4).

(C3) Simultaneous joint movements for force/power production

Simultaneous joint movements are used when producing forceful or powerful actions for a linked body segment chain that includes several of the major joints of the body. To ensure that this link system provides a firm base (see force principle F1) it is important that the muscle groups operate simultaneously. Therefore, forceful/powerful movements require muscles about joints to act synchronously. For example, when jumping for height, the hip and knee extensor muscles act simultaneously to generate high force; in the shot put, the drive from the trunk and legs occur simultaneously; in the bench press, the muscles activating the shoulder and elbow joint act together; in the sprint start, the hip, knee and ankle joint extend together.

Detailed biomechanical analysis has shown that in some cases, even though the muscles act synchronously, the joints do not extend synchronously. Typically in the vertical jump, the ankle joint extends after the hip and knee joints generate their main effort. As the ankle joint is weaker than these other two joints, its role is to maintain a firm base for the other stronger joints as they produce their effort. Once this diminishes, the ankle, which is kept in a position of flexion, extends making its contribution to the movement. The sequence of joint motion is a reflection of the relative strength and function of the joints rather than a violation of the principle.

(C4) Stretch–shorten cycle (SSC)

Many actions involve a pre-stretching of the muscles and tendons, which aids performance by enabling the highest muscle forces to be built up at the beginning of the movement. Consider a throw (or jump) in which the starting position is in the squat position. When the muscles begin to shorten, their force develops gradually. As this force develops, the movement velocity increases, as does the speed of shortening of the muscles, which are now less capable of generating force (as determined by the force-velocity relationship for muscle). The net effect is that the movement is less able to increase speed further. If the muscles are pre-stretched by a countermovement, the muscles are fully active at the beginning of the upward movement. This means that as the upward movement begins, they are generating their maximum effort. The muscle and tendon 'stretch' should be quickly followed by the 'shorten' in order to maximize this effect.

Specific performance (P) principles

Specific performance principles are a range of movement principles which identify the underlying factors relevant to specific aspects of performance.

(P1) Flight and projectile motion

An object which moves through the air under the influence of gravity is called a projectile. The key outcome of projectile motion is often the range, but sometimes the height reached and time of flight are important performance measures. The mechanical factors determining projectile motion are the height, angle and speed of release, with the speed of release being the most important. The effects of air resistance can be important in many situations, particularly those where the projectile is light, the relative velocity of the wind is high, or the object is shaped so as to have aerodynamic properties. The flight path of the projectile is modified accordingly and this as well as more complex effects, e.g. spin, need to be taken into account.

(P2) Speed–accuracy trade off

In the performance of many skills, the outcome is determined by both speed and accuracy. It is generally found that as the demands for accuracy increase, the speed of the movement decreases. For example, when kicking a football for accuracy, as in a penalty kick, a hard hit shot is less accurately placed than one hit less hard; in a basketball shot, the greater the distance of the throw, the less chance the ball will go into the basket; in the long jump approach, jumpers need to hit the take-off board accurately so there is a tendency to reduce their speed close to touchdown.

AN APPLICATION OF THE PHASE ANALYSIS MODEL AND MOVEMENT PRINCIPLES

A soccer penalty kick is used to illustrate how the phase analysis model and movement principles are used. The kick would typically be recorded on video and be available for repeat viewing, and inspection of individual frames. A series of still images have been extracted from such a video and presented in Figure 2.2, which cover the kick from take-off from the kicking leg on the last stride to follow-through.

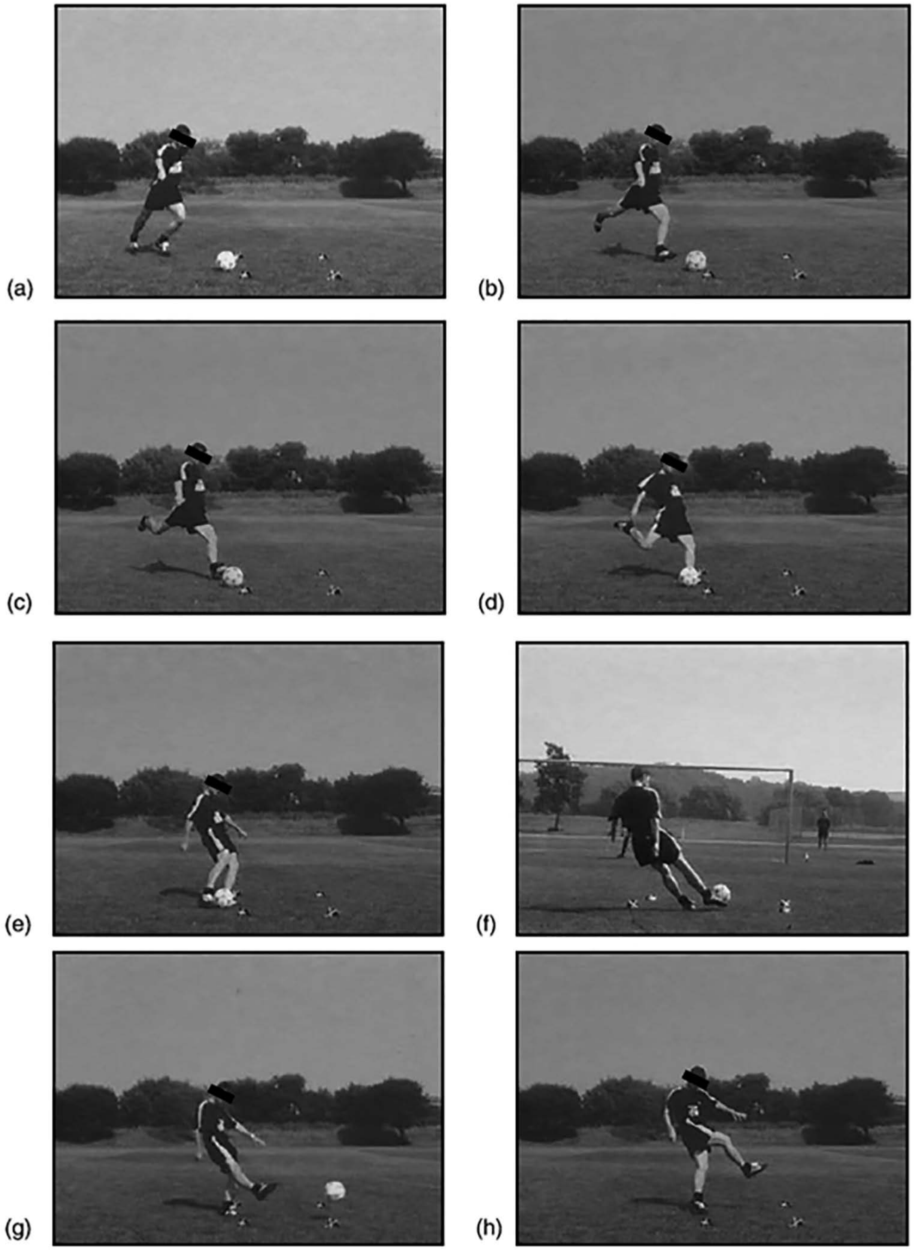
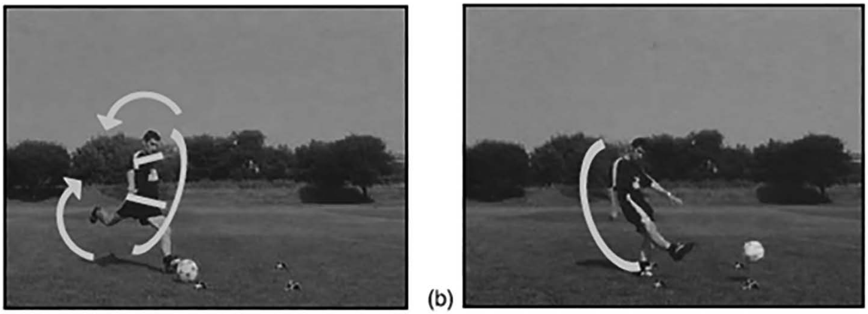


Figure 2.2 Still images of a penalty kick in soccer: (a) take-off from the kicking leg, (b) last stride, (c) touchdown support leg, (d) maximum knee flexion of kicking leg, (e) contact, (f) contact rear view, (g) post-impact, (h) follow-through.

In the phase analysis template (Table 2.5), each phase is identified along with the relevant sub-phases, key moments, critical features and movement principles. The movement principles can be the abbreviated names as used earlier. It can be seen from this analysis that several movement principles may apply at the

Table 2.5 Phase analysis model template for the soccer kick.

Phase description	*Preparation (run up)*			*Retraction (of kicking leg)*	*Action (swing of kicking leg)*		*Follow-through*	
Sub-phase description	1 Place ball and withdraw	2 Approach strides	3 (fig 2.2a) Last stride	1 SL placement of (2.2b) KL hip extension (2c)	1 (fig 2.2d) KL hip forward KL knee flexes	2 (fig 2.2e) KL knee extends	1 (fig 2.2g) KL knee full extension	2 (fig 2.2h) KL knee flexes
Key moments			KL take-off	KL max hip extension (2.2c)	KL max knee flexion	Impact	KL knee fully extended	KL knee flexed
Critical features			Stride length	KL hip extension (2.2c) Opposing arm back (2.2d)	KL max knee flexion	Body posture at impact	KL knee fully extended Opposing arm forwards	KL knee flexed
Movement Principles			ROM	ROM, SSC (stretch), AR	PDS	PDS, SSC (short-en), EPS	AR	



(a) (b)

Figure 2.3 Selected images from the soccer kick with indications of important movement principles. (a) Illustration of the stretch arc (stretch shorten cycle principle), retraction of the hips and the hip-shoulder separation (both ROM – range of motion principle) and the simultaneous retraction of the kicking leg and opposite arm (AR – action reaction principle). (b) Illustration of the shorten arc (SSC – stretch shorten cycle principle) in the follow-through.

same time. Some of these movement principles can be more easily appreciated by drawing appropriate indications on the images. For example, in Figure 2.3 the stretch shorten cycle, range of motion and action-reaction principles are easily appreciated from annotations to the images. Once a phase analysis model is completed, the coach is then able to view performances of the skill with a knowledge of what to look for in the movement (critical features) and how these relate to the mechanical performance of the skill (movement principles). For example, if a player needs to improve kick speed, then the coach would reasonably look at the range of motion achieved as indicated by the length of last stride, the degree of retraction of the hips and the use of the contralateral arm. Once these aspects of technique had been improved, the coach may then focus on the coordination of the movement, specifically the proximal-to-distal sequence of the movement and the general speed of execution which would improve the effect of the stretch–shorten cycle. Once these characteristics of technique had been developed, further improvement in performance may well come from the development of muscle strength characteristics. One should note that as strength changes so too may the technique, therefore a continual monitoring of the technique used should always be made.

AN APPLICATION OF THE HAY AND REID OUTCOME MODEL

The long jump is used to illustrate how the performance outcome model can be applied. The measure of performance in the long jump is the official distance jumped. As athletes must take off in front of the foul line, they often give themselves a margin of safety by taking off a few centimetres in front of the foul line. This distance is known as the toe-to-board distance. Thus, the actual distance jumped is almost always greater than the official distance. In fact, by simple addition, the actual distance jumped = the official distance + toe-to-board distance. This is an example of rule 1 for the construction of the model. The actual distance can then be divided into the take-off distance (the distance that the centre of mass is in front of the toes at take-off), flight distance (the distance the centre of mass moves through the air) and the landing distance (the distance that the centre of mass is behind the heels at landing). This is also an example of rule 1. Of these three distances, the most important is the flight distance. In flight, the body is a projectile and so the flight distance is governed by the mechanical variables that determine projectile flight. These are the height, speed and angle of projection at release (see performance principle P1). For practical reasons, it is more convenient to combine the speed and angle together and then use the velocity components in the vertical and horizontal directions. Thus, in this example the three projectile parameters used are height, horizontal velocity and vertical velocity of the centre of mass at release. This is an example of rule 2, which is based on mechanical relationships. A full hierarchical analysis for the long jump is given in Figure 2.4 (but also see Bartlett, 2007, for a more detailed development). Note that only the main factors are followed through. Nevertheless, the model is quite comprehensive and has been used by the author to provide scientific support for international long jumpers. The model identifies clearly what needs to be measured through biomechanical analysis. For example, the horizontal velocity of the centre of mass at touchdown and take-off are key factors. These can be measured from motion analysis during competition. This provides a strong rationale for the provision of biomechanical services during competitive events. If one wants to measure factors deeper into the hierarchy, such as force, then the model implies that a force platform needs to be used. So far, this has not been done in high-level competition and it may never be done.

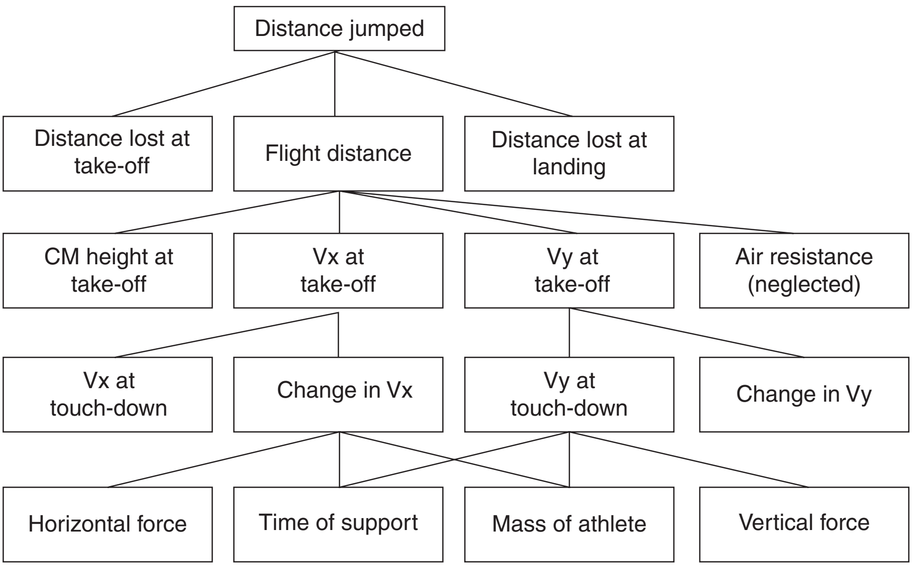


Figure 2.4 Performance outcome model for the long jump. (See text for further explanation.) CM, centre of mass; Vx, horizontal velocity; Vy, vertical velocity.

This type of information is therefore only gained in a training environment. The implication is that suitably instrumented training environments need to be available to provide comprehensive support to high-level athletes. Finally, we need to return to the issue of how this model can be used to analyse technique. Implicit in the model is the need to quantify the factors identified. This is the role of biomechanists but once this is done the results of their analyses can be used with the model to provide advice on aspects of technique. For example, it is clear in the model that performance is dependent on the horizontal velocity. Thus, it is also clear that maximizing this will be advantageous. It is also apparent from the model that vertical velocity is important, but the only place that vertical velocity can be generated is at take-off. Thus, attention should be focussed on the actions of take-off and the actions leading up to take-off. This then implies that a more detailed investigation of the technique during this phase is warranted, and that may be approached using the phase analysis model described in detail earlier.

RECENT THOUGHTS AND DEVELOPMENTS

The scope of qualitative analysis is greater than outlined earlier in the chapter. Texts on the topic (Knudson and Morrison, 2002; Bartlett, 2014) emphasise wider issues such as the preparation, observation, evaluation, diagnosis and intervention skills required to undertake a complete qualitative analysis with a purpose of improving 'performance'. This emphasises the origin of qualitative analysis which was in the coaching arena where its application was to aid the coaching process. Its formalisation, as noted earlier, is still relevant to this group, but is also appropriate to the professionalization of the sports biomechanists whose services are now sought by national governing bodies of sport in many countries. The reader is encouraged to consult the Knudson and Bartlett texts for a wider appreciation of the method.

While texts have focussed on the wider issues, researchers have attempted to solve other problems. One in particular which is worth noting is the concept of 'technical level'. Any skill may be performed badly or well and as beginners learn they tend to move from a poor to a good demonstration of the skill. The research issue is whether this progress can be determined using qualitative analysis. Marques-Bruna et al. (2007) attempted to determine the technical proficiency of 187 boys and girls aged 6–11 years, and 31 male and female adults in a stationary ball-kicking task. Phase analysis was used to divide the event into five relevant components, and for each, three levels of technical proficiency were defined and scored. For example the first component was the approach and was classified as straight (1 point), angled (2 points) and curved (3 points). The other components were opposite arm movement, foot placement, contact pattern and follow-through. The highest technical level was associated with a mature skilled kick and the greater technical proficiency was associated with a higher score. The results were clearly able to distinguish between ages and sexes, with older children demonstrating a higher technical proficiency than younger children, and boys showing a higher technical proficiency than girls. These differences were evident in each component as well as overall. Interestingly the adults were little better than the 11-year-old boys. This novel approach introduced a new concept into the literature and has shown it to be sensitive to skill development of children on a large scale. Such an analysis using a quantitative approach would be impossible.

Finally it is worth commenting on the advancement of computer technology. The introduction of low cost video analysis has enabled qualitative analysis to occur, but this is now several decades old. The introduction of low cost computing has enabled quantitative analysis to progress. So the question is whether low cost computing can be combined with image analysis to advance qualitative analysis? Certainly developments have taken place which allow drawings and other information to easily be superimposed onto images to produce diagram-pictures, as in Figure 2.3, which could easily be used in the field in real time to provide feedback to performers. Computing technology is advancing rapidly so perhaps the best advice is to be . . . observant!

CONCLUSION

This chapter has sought to introduce the methods relevant to the qualitative analysis of sports (and exercise) skills. Qualitative analysis is appropriate to the biomechanical analysis of technique and two major approaches have been identified for this. The phase analysis model is appealing in that it is based on a sequential breakdown of performance, performed from visual images gained, typically, from video analysis. By identifying phases and sub-phases of movement, the key moments and critical features, it is possible to identify the principles that govern performance. Using these principles, faults can be diagnosed from which it is hoped that performance may be improved. In order to use this model effectively, a sound knowledge of movement principles is needed. These are not readily available in the literature so the opportunity has been taken to outline them in this chapter. In contrast, the second approach is based on performance outcome. This leads to a hierarchical model of performance which identifies the relationships and mechanical factors that govern performance. It is implicit in this model that many of these factors need to be measured quantitatively. Based on this information it is possible to gain an insight into the key factors that influence performance and attend to these directly (such as the importance of approach speed in the long jump) or to use this information to guide a more 'technique' orientated analysis using the previous phase analysis model. The tools provided in this chapter should enable the enthusiastic student of sport and exercise science to undertake an effective analysis of technique using qualitative methods.

REFERENCES

Bartlett, R. (2007) *Introduction to Sports Biomechanics*, 2nd edn. London: Routledge.
Bartlett, R. M. (2014) *Introduction to Sports Biomechanics: Analysing Human Movement Patterns*, 3rd edn. London: Routledge.

Chow, J. W. and Knudson, D. V. (2011) 'Use of deterministic models in sports and exercise biomechanics research', *Sports Biomechanics*, 10: 219–233.
Hay, J. G. and Reid, G. (1982) *Anatomy, Mechanics and Human Motion*. Englewood Cliffs, NJ: Prentice-Hall.
Knudson, D. V. (2002) 'Qualitative biomechanical principles for application to coaching', *Sports Biomechanics*, 6: 109–118.
Knudson, D. V. and Morrison, C. S. (2002) *Qualitative Analysis of Human Movement*, 2nd edn. Champaign, IL: Human Kinetics.
Lees, A. (2002) 'Technique analysis in sports: a critical review', *Journal of Sports Sciences*, 20: 813–828.
Marqués-Bruna, P., Lees, A. and Grimshaw, P. (2007) 'Development of technique in soccer', *International Journal of Coaching Science*, 1(2): 51–62.

7 Diagnostic tests for regression

Introduction

After completing data collection and prior to beginning the formal regression analyses, it is incumbent on the investigator to screen their data. Typically, this process is called data cleaning in which the investigator examines each variable in the data set to assure that errors have not been made. Errors can range from simple typographic mistakes such as transcribing the subject's weight in pounds when in fact kilograms were the units needed, to errors which may not be initially or readily obvious and can potentially skew the results of your formal statistical analyses. Therefore, the present chapter will focus on this latter error by introducing established data screening procedures used to identify outliers within a given data set.

Review of regression assumptions

As discussed in the ANOVA section, quantitative models have assumptions which first need to be met in order to generate valid conclusions about the population of interest. Similarly, regression analysis has its own set of assumptions: (i) normality; (ii) linearity; (iii) homoscedasticity; (iv) multicollinearity; and (v) independence.

Normality. Normality refers the data for each variable having a normal distribution and not being skewed in one direction or another.

Linearity. Linearity, as named, refers to the linear relationship amongst the independent and dependent variables. Therefore, if data are curvilinear then any predictions from a linear model may be biased. Linearity may be detected through bivariate scatterplots and evaluating whether the plotted points create a straight-line association. The observed versus the predicted values may also be plotted, which will result in a scatterplot with the data points distributed around the regression line.

Homoscedasticity. Homoscedasticity indicates that the variation for the residuals is constant without noticeable patterns. That is, the variances of the dependent variable at the various levels of the independent variable are approximately equal. To determine if the homoscedasticity assumption has been violated (i.e., heteroscedasticity), plot the residuals versus the predicted values. The plotted values should be equally distributed along the zero point forming a rough rectangle. Violation of homoscedasticity would be indicated by the plotted values flaring in or out (> or <) like a sieve. Note that this assessment (plotting the residuals vs. predicted values) can also be used to assess multivariate normality and linearity.

Multicollinearity. Multicollinearity refers to extremely high correlations (usually 0.90 or above) between predictors, or between predictors and the criterion. Correlations of 0.90 or higher between variables can be problematic causing strange analytical results – for example, extreme standard errors thus producing Type II errors in significance of predictors. Variables correlated at 0.90 or higher may be deemed "the same" variable essentially, since the overlapping variance is at least 81 percent. That being said, research in the medical sciences and related fields may indeed find variables so highly correlated are important to include in the regression since medical and health outcomes are highly determined or influenced by biological markers. If the researcher deems multicollinearity to be problematic, consider dropping one of the variables. Assessment of multicollinearity is done by evaluating the bivariate correlation matrix between variables and criterion. Also, *tolerance* may be assessed, which is 1 minus the squared multiple correlation (SMC) of a predictor by the other predictors in the regression (1 – SMC). Extremely low tolerance values such as 0.10 can indicate high multicollinearity, although values <0.20 may also be of concern. The *variance inflation factor* (VIF) may also be assessed, calculated as 1/tolerance. Values greater than 10 suggest high multicollinearity, and if the average of the VIF values across variables exceeds 1.0, that is also indicative of problematic multicollinearity.

Independence. Independence refers to the independence of errors in the prediction of the criterion. Nonindependence of errors of prediction suggests there is some dependency between the order of cases in the data and the regression variables. One example of such a dependency bias would be "time," where due to interviewer effects those sampled earlier during the study period show more varied responses compared to those later sampled. Violation of this assumption can be assessed by plotting the residuals by the order cases were assessed. A formal test of this assumption is the *Durbin-Watson test*, which looks at the autocorrelation of errors by the case sequence. The Durbin-Watson test ranges between 0 and 4, with values close to 2 indicating uncorrelated errors (values close to 0 or 4 indicate extreme positive or negative autocorrelation, respectively). Values between 1.5 and 2.5 are indicative of independence.

Identifying outliers

It is essential prior to performing your formal statistical analyses to screen your data. This is important because it will assist in identifying case(s) that are outliers. An outlier may be a data point with an extremely higher or lower value related to other data points for that variable. Outliers may be simple typographical errors (such as entering the number 500 when you meant to enter 50) or more complex. Nevertheless, it is important to remember that outliers may potentially distort the results of your formal test statistic. A number of different data screening techniques are used to identify outliers.

Bivariate plots. Bivariate scatterplots of the predictors and the criterion can reveal cases that across pairs of variables have extreme high or low values. Cases that fall outside the "swarm" of values in these plots may be outliers and should be evaluated further for adjustment or exclusion. It may be assessed through bivariate plots of the predictors, and is revealed by looking for cases that are removed from the swarm of plotted values. As a general example, say two variables (age and income) are used to predict a measure of community status. Focusing on age and income, someone who is 18 years of age might be typical in a community sample, as would someone making $100,000 per year. A bivariate plot of these data might reveal an 18 year old who earns $100,000 per year – such a case may be considered high in leverage if it appears removed or separate from the other plotted values.

Standardized residuals (ZRESID). The ZRESID are expressed on the z-score scale and are calculated by dividing the residual ($Y_i - Y'$) by the standard deviation of the sample deviations. A large standardized residual indicates a case is ill-fit by the regression line and may be an outlier.

Adjusted predicted value (ADJPRED). The ADJPRED is the adjusted predicted value, where the predicted value from the regression equation is calculated removing the *i*th case ($Y' = a + b_i^{(i)} x_i$), where (i) indicates the removal of the *i*th case. If the ADJPRED is noticeably different from the original predicted value, the *i*th case may be an outlier.

Deleted residuals (DRESID) and studentized deleted residuals (SDRESID). The DRESID is the deleted residual, which is a recalculation of the residuals with the *i*th case removed ($Y_i - Y'^{(i)}{}_i$). If the DRESID is noticeably different from the calculated residual, then the case may be overly influential. The SDRESID is the studentized deleted residual, which is calculated by taking the deleted residual and dividing it by its standard error. As with the DRESID, if the SDRESID is noticeably different from the SRESID (the studentized residual), then the case may be overly influential.

Influential data point

Leverage. Leverage reflects a case that may be considered distant from other cases across the predictors or independent variables. Leverage is not inclusive of the dependent variable, but instead focuses on the independent variables or predictors. Formal calculations to assess leverage are offered in various statistical programs, with leverage values ranging from 0 to 1. The resulting values may be compared to a cutoff value – if a leverage value exceeds the cutoff, the case may have undue influence on the resulting regression line. The cutoff value is calculated as follows (with *k* indicating the number of predictors in the regression model):

$$2\left(\frac{k+1}{N}\right)$$ **Formula 10.1 Leverage cutoff value**

Cook's distance (Cook's D). Influence and its resulting assessment using Cook's distance (Cook's D) also reflects a case's influence on the regression line. It is a measure of change in the residuals when the *i*th case is removed. Although one can simply note the difference between the residual and the deleted residual for a given case to provide an assessment of a case's influence, Cook's D provides a broader look since it encompasses change in all of the residuals. Larger values of Cook's D indicate greater case influence. A rough cutoff to indicate the case may be problematic can be calculated using $4/(N-2)$, with N being the total number of cases. If Cook's D for a case exceeds the cutoff, it may have an undue influence on the regression line.

$$C_i = \frac{\sum (Y'^{(i)} - Y')^2}{(p+1)S^2}$$ **Formula 10.2 Cook's distance**

Research question

The investigator is interested in performing diagnostics on the regression model they are planning to develop. Below we present the interpretations of the various methods of screening the data set. Thereafter, we present the pulldown menu instructions and syntax commands for achieving each analyses.

Interpretation of data screening

Normality, linearity, and homoscedasticity

In starting to make our assessment, we begin by assessing *normality*. This is done by evaluating histograms of the variables in the planned regression, and is considered a univariate or single variable assessment (a multivariate normality assessment will be covered shortly). Figures 7.1 and 7.2 illustrate the syntax and resulting histogram for two of the variables (oxygen uptake and age). Keep in mind that the criterion variable is oxygen uptake and will be referred to as such in the examples below. For both histograms, a slight skew is noted, but in general it may be concluded that these variables are approximately normal.

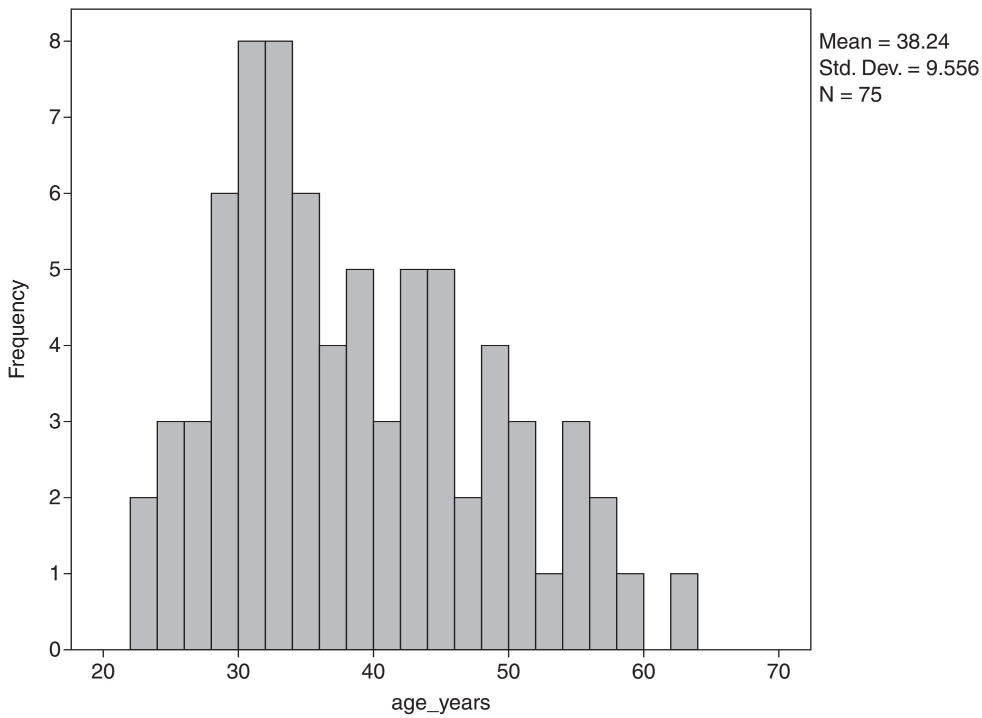


Figure 7.1 Histogram for the age variable, reprint courtesy of International Business Machines Corporation.

Source: © International Business Machines Corporation.

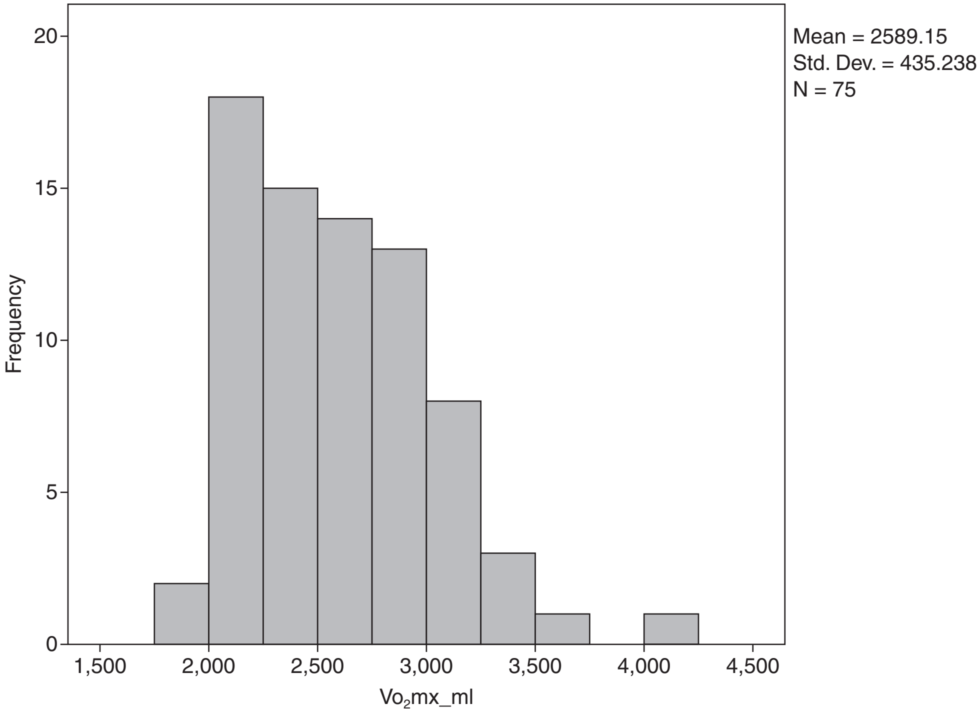


Figure 7.2 Histogram for $\dot{V}o_2$ *max* variable, reprint courtesy of International Business Machines Corporation.

Source: © International Business Machines Corporation.

Next, we assess *linearity* using bivariate scatterplots which is performed for all variables in the planned regression. In Figure 7.3, we present syntax for the bivariate plots of the criterion with age, and with weight. In these graphs, we wish to see straight-line associations (an absence of curvilinearity). Although not perfectly aligned, the graphs suggest linearity is tenable. Since these are bivariate assessments, a multivariate assessment should also be performed (to be demonstrated in the next section).

Homoscedasticity is assessed by plotting the standardized predicted values by standardized residuals. This same multivariate plot is also used to assess multivariate normality and linearity. To assess homoscedasticity (and multivariate normality and linearity), run the planned regression analysis and request a plot of the standardized predicted values by the standardized residuals. This is demonstrated in Figure 7.4. To meet all three multivariate assumptions, the plotted values should be equally distributed along the zero point forming a rough rectangle. Violation of homoscedasticity would be indicated by the plotted values flaring in or out (> or <) like a sieve. Violation of normality would be indicated by the plotted cases clustered either below or above the zero mark (as opposed to being equally distributed). Violation of linearity would be indicated by the plotted cases forming a ∪ or ∩ association thereby suggesting a curved association. In Figure 7.4, the cases meet the assumptions of

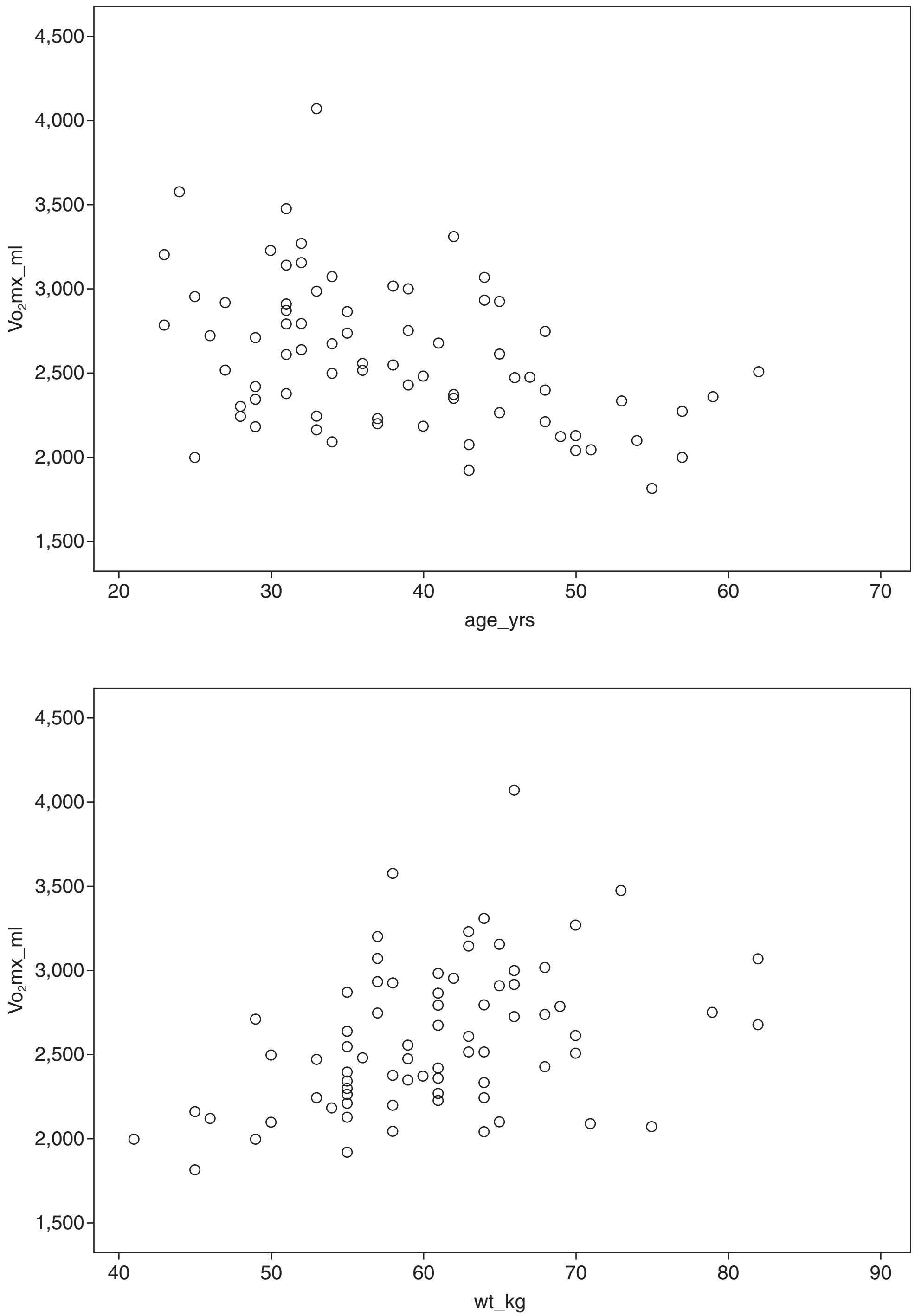


Figure 7.3 Bivariate plots of the criterion with age and weight variables, reprint courtesy of International Business Machines Corporation.

Source: © International Business Machines Corporation.

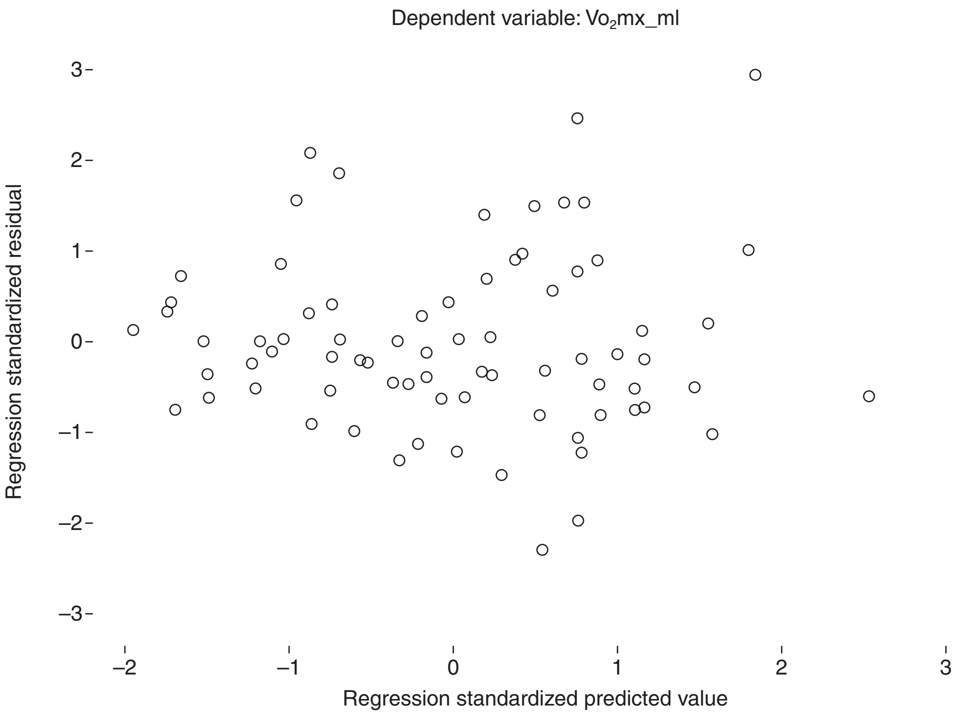


Figure 7.4 Homoscedasticity plot scatterplot, reprint courtesy of International Business Machines Corporation.

Source: © International Business machines corporation.

multivariate normality and linearity. However, there may be an issue with homoscedasticity since the plotted cases flare outward (<), which may produce a degraded solution and thus increase the likelihood of Type II error.

The graph above (Figure 7.4) illustrates a fairly good spread of the standardized residuals plotted against standardized predicted values. If we draw a horizontal line at the zero point along the ordinal axis (regression standardized residual), the plot points equally fall above/below the line, suggesting multivariate normality. Using this same horizontal line, we may also assess homoscedasticity. Here, the plot points generally fall above/below the line, although there is a slight flaring pattern suggesting heteroscedasticity, which may produce a degraded solution. Thus, borderline regression results should be interpreted carefully. If we next draw a line of best fit across the plot points, there appears to be a straight-line association (no clear curvilinear association), suggesting multivariate linearity.

Multicollinearity is assessed using bivariate correlations, tolerance, and the variance inflation factor (VIF). Bivariate correlations are presented in Table 7.1, with tolerance and the VIF presented in Table 7.2. None of the bivariate correlations across variables exceed 0.90. The tolerance values are well above 0.10, and the VIF values well below 10 (with the average VIF exceeding 1.0 as well). Extreme multicollinearity does not appear to be an issue within our data.

Table 7.1 Correlations, reprint courtesy of International Business Machines Corporation

Correlations

		Vo_2mx_ml	*wt_kg*	*ht_meter*	*age_yrs*	*Times/wk_exercise*	*Intensity of subjects training (6–20)*
Vo_2mx_ml	**Pearson Correlation**	1	.413	.577	–.442	.216	.373
	Sig. (2-tailed)		.000	.000	.000	.063	.001
	N	75	75	75	75	75	75
wt_kg	**Pearson Correlation**	.413	1	.599	–.058	.069	–.197
	Sig. (2-tailed)	.000		.000	.624	.554	.091
	N	75	75	75	75	75	75
ht_meter	**Pearson Correlation**	.577	.599	2	–.331	.144	.062
	Sig. (2-tailed)	.000	.000		.004	–.216	.594
	N	75	75	75	75	75	75
age_yrs	**Pearson Correlation**	–.442	–.058	–.331	.1	–.134	–.139
	Sig. (2-tailed)	.000	.624	.004		.252	.236
	N	75	75	75	75	75	75
Times/wk_exercise	**Pearson Correlation**	.216	.069	.144	–.134	1	.206
	Sig. (2-tailed)	.063	.554	.216	.252		.077
	N	75	75	75	75	75	75
Intensity of subjects training (6–20)	**Pearson Correlation**	.373	–.197	.062	–.139	.206	1
	Sig. (2-tailed)	.001	.091	.594	.236	.077	
	N	75	75	75	75	75	75

Table 7.2 Excluded variables, reprint courtesy of International Business Machines Corporation

Excluded Variables[a]

Model		*Beta in*	*t*	*Sig.*	*Partial correlation*	*Colinearity statistics*		
						Tolerance	*VIF*	*Minimum tolerance*
1	**Times/wk_exercise**	.112[b]	1.226	.224	.145	.971	1.030	.549
	Intensity of subjects training (6–20)	.377[b]	4.477	.000	.472	.905	1.105	.533

Source: © International Business Machines Corporation.

Notes

a Dependent Variable: Vo_2mx_ml.

b Predictors in the Model: (Constant), age_yrs, wt_kg, ht_meter.

Multicollinearity is assessed in the table above (Table 7.2) focusing on the tolerance and VIF (variance inflation factor). The two statistics are related (1/VIF = tolerance). Low tolerance values (<0.20) suggest possible multicollinearity. For VIF, a single value greater than 10 is problematic, while an "average" VIF across all variables exceeding 1.0 may suggest multicollinearity issues. Evaluating Model 1 above, no variables have tolerance values less than 0.20. For VIF, all values are well below 10, and the average VIF exceeds 1.0. Therefore, there does not appear to be multicollinearity in our model that may bias the regressions results.

Independence of errors of prediction

This is assessed requesting the Durbin-Watson statistic, which tests the assumption of independence of errors of prediction. Table 7.3 contains the necessary syntax to produce the Durbin-Watson statistic. As noted earlier, values between 1.5 and 2.5 are indicative of independence. In our example the Durbin-Watson statistic is 2.082, well within the acceptable range.

As noted earlier, the Durbin-Watson test assesses autocorrelation of errors by the case sequence to evaluate independence of errors. That is, are adjacent case residuals correlated, or independent? As shown in Table 7.3, our value of 2.082 is close to 2.0 (indicative of independence).

Table 7.3 Model summary, reprint courtesy of International Business Machines Corporation

Model Summary[c]

Model	*R*	*R square*	*Adjusted R square*	*Std. error of the estimate*	*Durbin-Watson*
1	.650[a]	.422	.398	337.741	
2	.743[b]	.553	.520	301.511	2.082

Source: © International Business Machines Corporation.

Notes

a Predictors: (Constant), age_yrs, wt_kg, ht_meter.

b Predictors: (Constant), age_yrs, wt_kg, ht_meter, Times/wk_exercise, intensity of subjects training (6–20).

c Dependent Variable: Vo_2mx_ml.

Outlier assessment

Outliers are assessed through a number of different ways. Bivariate plots (for example, those in Figure 7.2) can be assessed for data points beyond the swarm of plotted points. In Figure 7.3, most cases are reasonable, although a few straggling cases are noted in each plot. The question to ask, however, is whether these straggling cases are indeed outliers and whether they are having an undue influence on the overall regression? Further assessment using additional techniques will help answer this question.

The standardized residuals are also evaluated to assess outliers, along with the adjusted predicted values and deleted residuals. Table 7.4 is the syntax and output to assess the standardized residuals. The most extreme ten cases are printed, and a histogram of the residuals is also presented. Overall, two cases have standardized residuals exceeding ± 2.56 indicating problematic fit. Whether these cases have an undue influence on the regression line will be assessed in a later section, but here we should be at least moderately concerned that these cases may be problematic. The histogram for the standardized residuals looks appropriate.

In the table above (Table 7.4), the studentized deleted residual is provided as a means of determining overly influential cases in our data set. Those cases exceeding a standardized value of 2.56 ($p < 0.01$) are cases with undue influence. Here, two cases (#28 and #57) exceed the 2.56 cutoff, and should be further evaluated. Note the case number is not the subject number in the data set, but literally the "row" number associated with the case.

The adjusted predicted value may also be evaluated for outliers by plotting it by the predicted value. Figure 7.5 presents the syntax and resulting plot. Cases that fall off the diagonal are indicative of possible outlying cases. In Figure 7.5, it appears most cases are in alignment.

The deleted residuals and studentized deleted residuals may also be evaluated for outliers. This is done by plotting the deleted residual by the calculated residual, and

Table 7.4 Outlier statistics, reprint courtesy of International Business Machines Corporation

Outlier Statistics[a]

		Case number	*Statistics*
Stud. Deleted Residual	**1**	28	3.346
	2	57	2.644
	3	73	–2.478
	4	55	2.172
	5	18	–2.089
	6	48	1.929
	7	12	1.616
	8	16	1.607
	9	35	1.590
	10	7	1.566

Source: © International Business Machines Corporation.

Note

a Dependent Variable: Vo_2mx_ml.

plotting the studentized deleted residual by the studentized residual. Cases that fall away from the diagonal are indicative of possible outlying cases. For both graphs (Figure 7.6), there is no evidence of outlying cases.

Influential data points

Influential data points may be examined using a centered leverage measure and Cook's distance (Cook's D). As shown in Table 7.5, we request a case-wise plot of

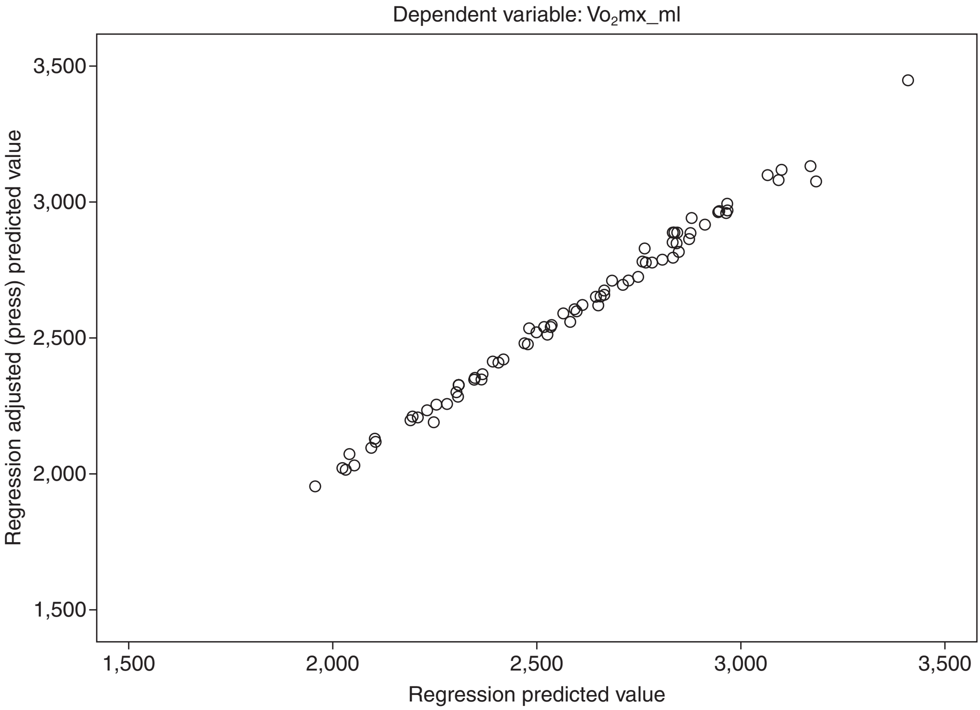


Figure 7.5 Scatterplot, reprint courtesy of International Business Machines Corporation.

Table 7.5 Casewise diagnostics, reprint courtesy of International Business Machines Corporation

Case number	*Std. residual*	*Vo_2mx_ml*	*Residual*	*Centered leverage value*	*Cook's distance*
28	2.947	4073	888.672	.096	.200
55	2.080	2934	627.136	.020	.026
57	2.468	3578	744.055	.040	.061
73	–2.287	2074	–689.569	.072	.088

Note
a Dependent Variable: Vo_2mx_ml.

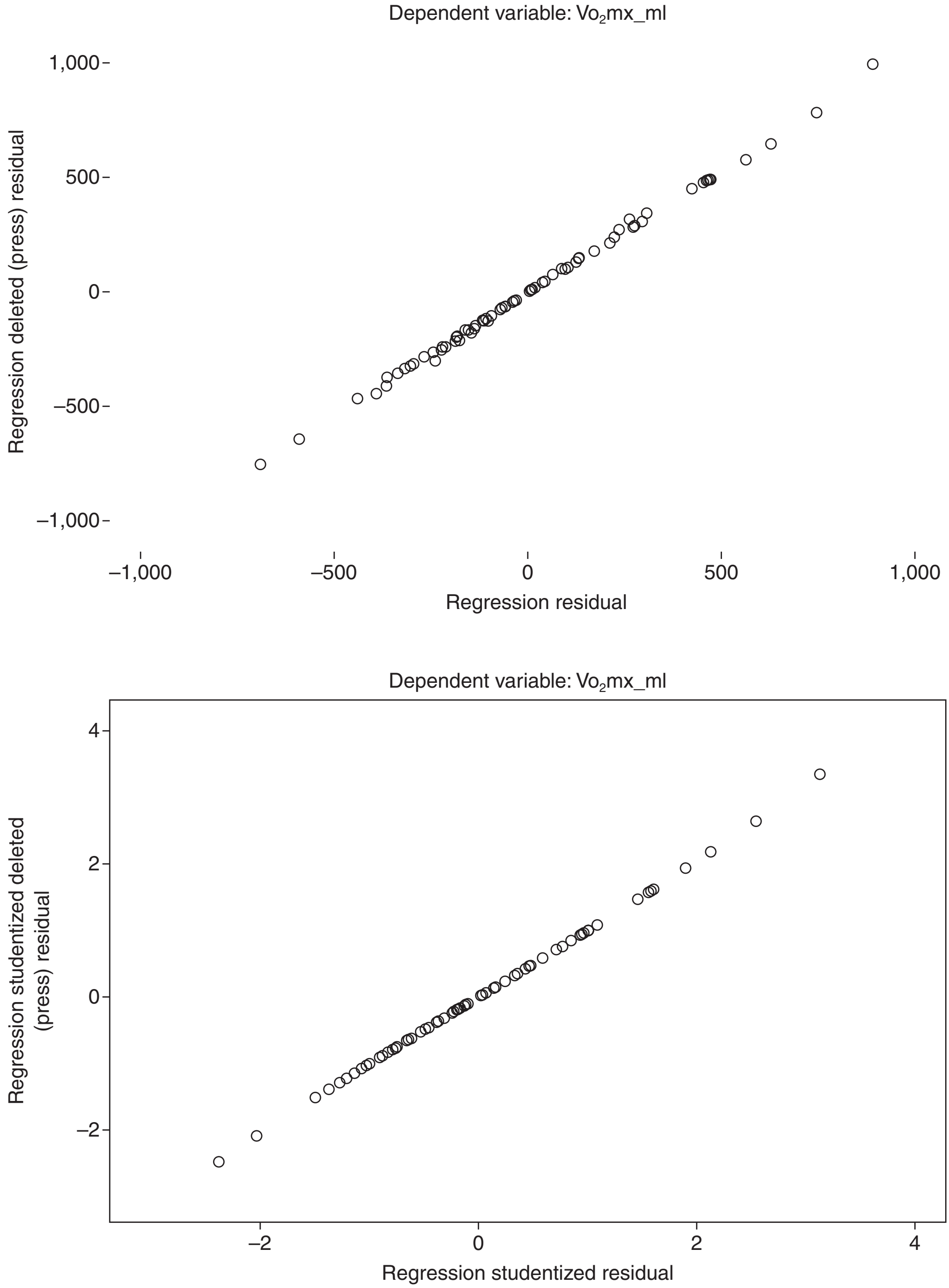


Figure 7.6 Scatterplot, reprint courtesy of International Business Machines Corporation.
Source: © International Business Machines Corporation.

cases with residuals greater than 2 using the "OUTLIERS(2)" command, which selects only those cases with standardized residuals ± 2. The residual (RESID) and standardized residual (ZRESID) is also presented for comparison purposes.

Using Formula 10.1 from earlier in the chapter, we calculate a leverage cutoff of 0.16. Below, k is the number of predictors in the regression, and N is the number of total cases. The values for leverage in Table 7.5 are all below the calculated cutoff of 0.16.

Leverage cutoff calculation

$$2\left(\frac{k+1}{N}\right) \tag{1}$$

$$2\left(\frac{5+1}{N}\right) \tag{2}$$

$$2\left(\frac{6}{75}\right)=0.16 \tag{2}$$

For Cook's distance, we calculate a cutoff using $4/(N-2)$, with N being the total number of cases. For our example, we calculate a cutoff of 0.0548: $4/73=0.0548$. Next, we evaluate the extreme cases in the case-wise table from Table 7.5 – three of our cases violate the cutoff, and therefore may have an undue influence on the overall regression results.

Sample write-up

Typically, there is no write-up for the diagnostics analyses. These are performed prior to the formal regression analyses. Nevertheless, it is important to conduct these diagnostics as discussed above.

Using SPSS pulldown menu to generate histogram graph

1. Click **Graphs**, then move cursor over **Legacy Dialogs**, then over cursor over **Histogram** then left click.
2. Click **age_yrs**, then move to **Variable** box.
3. Then click **OK**.
4. Follow steps 1–3 to generate histogram for the oxygen uptake (Vo_2mx_ml) variable.

Syntax to generate histogram graph

```
GRAPH
/HISTOGRAM=age_yrs.
GRAPH
/HISTOGRAM=Vo2mx_ml.
```

Using SPSS pulldown menu to generate bivariate scatter plot

1. Click **Graphs**, then move cursor over **Legacy Dialogs**, then over cursor over **Scatter/Dot** then left click.
2. Click **Simple Scatter** and then click **Define**.
3. Click **age_yrs** and move to X-axis box.
4. Click **Vo$_2$mx_ml** and move to Y-axis box.
5. Click **OK**.
6. Follow steps 1–5 to generate bivariate scatterplot for weight and oxygen uptake.

Syntax to generate bivariate scatterplots

```
GRAPH
/SCATTERPLOT(BIVAR)=age_yrs WITH Vo2mx_ml
/MISSING=LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR)=wt_kg WITH Vo2mx_ml
/MISSING=LISTWISE.
```

Using SPSS pulldown menu for multivariate normality, linearity, and homoscedasticity

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.
2. Click **Vo$_2$max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Plots**, then click ***ZPRED**, and move to “X:” box.
11. Click **Plots**, then click ***ZRESID**, and move to “Y:” box.
12. Click **Continue**.
13. Click **OK**.

Syntax to generate multivariate normality, linearity, and homoscedasticity

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10) CIN(95)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/SCATTERPLOT=(*zresid,*zpred).
```

Using SPSS pulldown menu for bivariate correlations

1. Click **Analyze**, then move cursor over **Correlate** and then move cursor over **Bivariate** and click.
2. Click **Vo_2mx_ml**, then move to Variables box.
3. Repeat step #2 for wt_kg, ht_meter, age_yrs, timeperw, and intensity.
4. Note that Pearson is checked and two-tailed is also checked.
5. Flag significant correlations box should be checked.
6. Click **OK**.

Syntax to generate bivariate correlations

```
CORRELATION
/VARIABLES Vo2mx_ml wt_kg ht_meter age_yrs timeperw intensty.
```

Using SPSS pulldown menu for tolerance and VIF

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.
2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Statistics**, then check the box for the following:
    a. Collinearity diagnostics.
    b. Confidence intervals level (%): 95.
11. Click **Continue**.
12. Click **OK**.

Syntax to generate tolerance and VIF

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty.
```

Using SPSS Pulldown Menu for Independence of Errors (Durbin-Watson Test)

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.

2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Statistics**, then check the box for the following:

    a. Collinearity diagnostics.
    b. Confidence intervals level (%): 95.
    c. Durbin-Watson.

11. Click **Continue**.
12. Click **OK**.

Syntax to generate independence of errors (Durbin-Watson test)

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/RESIDUALS DURBIN.
```

Using SPSS pulldown menu standardized residuals

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.
2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Paste**.

    a. This will bring up the syntax window with your syntax for the above steps.
    b. Then type in the following syntax on the last line:

        i. /residuals=histogram(sdresid) outliers(sdresid).

11. Then click the green play (▶) icon on the top of the menu.

Syntax to generate standardized residuals

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10) CIN(95)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/residuals=histogram(sdresid) outliers(sdresid).
```

Using SPSS pulldown menu for adjusted predicted values

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.
2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Paste**.
    a. This will bring up the syntax window with your syntax for the above steps.
    b. Then type in the following syntax on the last line:
        i. /SCATTERPLOT=(*adjpred, *pred).
11. Then click the green play (▶) icon on the top of the menu.

Syntax to generate adjusted predicted values

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10) CIN(95)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/SCATTERPLOT=(*adjpred, *pred).
```

Using SPSS pulldown menu for deleted residuals and studentized deleted residuals

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.

2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Paste**.

    a. This will bring up the syntax window with your syntax for the above steps.
    b. Then type in the following syntax on the last line:

        i. /SCATTERPLOT=(*DRESID, *RESID) (*SDRESID, *SRESID).

11. Then click the green play (▶) icon on the top of the menu.

Syntax to generate deleted residuals and studentized deleted residuals

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10) CIN(95)
/NOORIGIN
/DEPENDENT Vo_2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/SCATTERPLOT=(*DRESID, *RESID) (*SDRESID, *SRESID).

Using SPSS pulldown menu for leverage and Cook's D

1. Click **Analyze**, then move cursor over **Regression** and then move cursor over **Linear** and click.
2. Click **Vo_2max_ml**, then move to Dependent box.
3. Click **wt_kg**, then move to Independent(s) box.
4. Click **ht_meter**, then move to Independent(s) box.
5. Click **age_yrs**, then move to Independent(s) box.
6. Click **Next**.
7. Click **timeperw**, then move to Independent(s) box.
8. Click **intensity**, then move to Independent(s) box.
9. You should now be on Block 2 of 2.
10. Click **Paste**.

    a. This will bring up the syntax window with your syntax for the above steps.
    b. Then type in the following syntax on the last line:

        i. /casewise=resid zresid cook lever outliers(2).

11. Then click the green play (▶) icon on the top of the menu.

Syntax to generate for leverage and Cook's D

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10) CIN(95)
/NOORIGIN
/DEPENDENT Vo2mx_ml
/METHOD=ENTER wt_kg ht_meter age_yrs
/METHOD=ENTER timeperw intensty
/casewise=resid zresid cook lever outliers(2).
```

Chapter 5

Physical activity for mental health

Naomi Ellis and Leon Meek

Mental health

The term 'mental health' has been defined as the emotional and spiritual resilience enabling people to enjoy life and cope with adversity (Faulkner and Taylor 2005). While mental health is used positively to indicate a state of psychological wellbeing (Pilgrim 2005), the term 'mental health problem' refers to a negative aspect indicating a disruption to how people think and feel (Mentality 2004). The terms 'mental health problems' and 'mental illness' are often used interchangeably. Yet, references made to 'mental health problems', rather than specific diagnoses (e.g. schizophrenia) or 'mental illness', are thought to be less stigmatising and reduce negative connotations (Pilgrim 2005). In this chapter 'mental health problems' is used as an umbrella term, but 'mental illness' is sometimes used when discussing specific clinical diagnoses.

Together, mental, neurological, and substance-use disorders account for 13 per cent of the total global burden of disease, with depression as the single largest cause of disability worldwide (World Health Organisation 2013). In the UK, 22.8 per cent of the total burden of disease was attributable to mental health problems (WHO 2008), with one in four adults experiencing them during their lifetime (Singleton et al. 2001). However, some studies suggest that the lifetime prevalence is higher (e.g. Kessler et al. 2005). Furthermore, mental health affects, and is affected by, other non-communicable diseases, such as cancer and cardiovascular disease (Royal College of Psychiatrists 2010), which also have common risk factors (e.g. low socio-economic status, alcohol use, and stress).

The scale of this disease carries a considerable economic cost. In England, the annual cost of poor mental health in 2009–10 was estimated at £105 billion, including health and social care, lost economic output, and human suffering (Centre for Mental Health 2011). And the problem is increasing: the 2009–10 figures marked a 36 per cent increase from 2002–03 (Centre for Mental Health 2011), with projections of a further 14 per cent increase in the number of people in England experiencing mental health problems by 2026 (McCrone et al. 2008). The cost of this growing prevalence is reflected most in health and social care costs, which rose from £12.5 billion to £21.3 billion between 2002–03 and 2009–10 (a 70 per cent rise; Centre for Mental Health 2011). Mental health problems account for almost a quarter of all primary care consultations (Department of Health 2004) and one third of all GP time (Social Exclusion Unit, 2004). This is compounded by the greater risk of comorbidities in this population group. People with mental health problems have a higher risk of physical health issues (WHO 2013), chronic disease (Brown et al. 2000; Connolly and Kelly 2005) and premature mortality (Harris and Barraclough 1998), placing even greater pressure on health services (Ansseau et al. 2004; The MaGPIe Research Group 2003).

In the context of a disproportionate need for health care, there is evidence that provision for those with mental health problems is inconsistent. Issues have been reported around access to services, effectiveness of interventions, poor holistic care, and a lack of attention to physical health needs (Mental Health Foundation 2013a). Traditionally, mental health problems have been treated through prescribed medication (Sin and Gamble 2003). However, the potential side effects (Haslam et al. 2004), low rates of adherence (Chue and Kovacs 2003), and subsequent relapse, are all concerns (Cramer and Rosenbeck 1998). This, combined with the cost of pharmacological treatments, prompted interest in alternative or adjunct therapies to help reduce the mental health burden (Lam and Kennedy 2004). Cognitive behavioural therapy (Haddad 2005) and other forms of psychotherapy are now offered to patients, although access is variable (Social Exclusion Unit, 2004). Ultimately, the rising cost of pharmacological and psychological interventions means that health services may soon be unable to meet the demands (Martinsen, 2000). Physical activity has come to the attention of researchers and practitioners alike as an alternative or adjunct therapy for mental health problems, because of wide ranging benefits for mental and physical health without the potential side effects of medication (Daley 2002; Priest 2007).

Physical activity and mental health

Physical inactivity is one of the leading risk factors for global mortality (World Health Organization 2010), with sedentary behaviour now recognised as a major public health issue (Health and Social Care Information Centre 2012). The benefits of being physically active are well documented (Biddle and Mutrie 2008) and include the prevention and management of mental health conditions (World Health Organization 2010). Despite awareness of the benefits of active lifestyles, a large proportion of the UK population are inactive. Only 56 per cent of adults are thought to meet the Chief Medical Officer's guidelines (HM Government 2014; Health and Social Care Information Centre 2012), although this is a likely over-estimate, as it is based on self-reported physical activity measurements. True levels of participation are likely to be much lower, as indicated by objectively measured physical activity data, which suggest that only 6 per cent of men and 4 per cent of women meet Government guidelines (NHS Information Centre for Health and Social Care 2009).

The economic cost of this widespread physical inactivity is considerable (British Heart Foundation National Centre 2013). Direct costs in the UK are estimated to be over £1 billion (Allender et al. 2007; Scarborough et al. 2011), which increases significantly when loss of productivity (in England £5.5 billion) and premature death of working age people (in England £1 billion) are taken into account (Ossa and Hutton 2002). Given the consistent evidence linking active lifestyles to health, increasing population physical activity levels would help to improve physical and mental health, and reduce the associated cost of hypokinetic chronic diseases (Department of Health 2011).

Physical activity has been shown to have both preventative and therapeutic benefits for mental health (Chen and Millar 1999; Walsh, 2011). Physically active individuals tend to report lower depression and symptoms of anxiety when compared with inactive participants (Cassidy et al. 2004; Hassmen et al. 2000; Pasco 2011). People whose activity levels fall below levels recommended for health benefit increase their risk of mental health problems (Lampinen et al. 2000). The reverse has also been observed; previously sedentary individuals can reduce their risk of mental health problems by taking up physical activity (Brown et al. 2005; Camacho et al. 1991; Chen and Millar 1999).

Physical activity for people with mental health problems

The mental health benefits of physical activity in the general population provide a rationale for therapeutic application in those with established mental health problems. Cross-sectional evidence linking physical activity and the risk of mental health problems is consistent (Harvey et al. 2010). Although there is not as much evidence of the cause-effect relationship (Morgan et al. 2013), it is broadly supportive (Ten Have et al. 2011; Strawbridge et al. 2002).

When we look to trials and intervention studies of exercise in clinical mental health populations, the therapeutic potential is apparent (Cooney et al. 2013; Ellis et al. 2007; Holley et al. 2011; Wolff et al. 2011). Often, researchers have focused on physical activity for reducing symptoms of depression and anxiety (Craft and Landers 1998; Craft and Perna 2004; Crone and Guy 2008; Hutchinson 2005; Pelham and Campagna 1991; Pelham et al. 1993; Priest 2007), but increases in self-confidence (Mind 2007), mood (Ellis et al. 2013; McDevitt et al. 2005; Randall et al. 2014), enjoyment (Fogarty and Happell 2005), and motivation (O'Kane and McKenna 2002), as well as reductions in positive symptoms of psychosis (Faulkner and Sparkes 1999), and stress (Carless and Douglas 2004).

In relation to some of the most prevalent or severe mental health problems, again, the evidence is supportive. For example, depression is a common condition, affecting 121 million people globally, and physical activity has been shown to help prevent its onset (Mammen and Faulkner 2013). A short ten-day aerobic exercise intervention targeting in- and out-patients diagnosed with major depressive disorder was associated with a decrease in depression (Dimeo et al. 2001). Hutchinson (2005) evaluated a twenty-week (which was long-term in comparison to most interventions) gym-based exercise intervention consisting of three 45-minute sessions per week. Follow-up data from 40 weeks identified a significant improvement, not only in cardiovascular fitness, but also in depression and participants' quality of life. There is also review-level evidence that physical activity can improve depression. A small but significant beneficial effect has been reported in children and adolescents (Brown et al. 2013), and in adults (Mead et al. 2008), although the latter review reported methodological limitations in many cases.

Another example is the use of exercise for the prevention and treatment of cognitive impairment and dementias. There is growing evidence that physical activity is a key modifiable risk factor that could be used to reduce the prevalence of Alzheimer's disease, as well as related comorbidities such as diabetes (Mehlig et al. 2014; Norton et al. 2014). A meta-analysis identified that action on seven modifiable risk factors could prevent around one third of global cases of Alzheimer's diseases, and that physical inactivity was the most potent factor (Norton et al. 2014). Indeed, there is growing review-level evidence that interventions that combine exercise with cognitive training can improve cognitive function, thereby helping to delay the age-related decline in mental ability, and reduce the prevalence of mild cognitive impairment and dementias (Law et al. 2014). Similarly, a review of exercise training for older people with cognitive impairment and dementia reported significant benefits for strength, fitness, functional and cognitive performance, and behaviour (Heyn et al. 2004). A later Cochrane review of exercise as a treatment for dementia did report benefits for the ability to perform activities of daily living and, possibly, for improving cognition, but also highlighted variation between results of trials warranting a degree of caution (Forbes et al. 2013).

Psychosis is an important group of severe mental health disorders, which encompasses schizophrenia, schizophreniform disorder, schizoaffective disorder, bipolar disorder, and major depression with psychotic features (Ehmann and Hanson 2002). Positive symptoms of such conditions, such as hallucinations and delusions, are particularly debilitating (American Psychiatric Association 1994). A 2007 review of exercise as an adjunct therapy for psychosis was generally supportive of a positive effect, but stated the need for more consistency to allow the benefits to be quantified and to identify the most successful types of intervention (Ellis et al. 2007). Subsequently, a review of RCTs supported the use of physical activity as part of a multidisciplinary treatment of schizophrenia (Vancampfort et al., 2012), with more recent calls for modern multimodal intervention approaches that include physical activity to improve psychopathology and cognitive symptoms in psychosis (Malchow et al. 2013).

As noted above, some researchers express a need for caution when interpreting evidence in this area. Where discrepancies in the strength and consistency of relationships are observed, there are various considerations, such as the small number of RCTs, typically small samples, short follow-up periods, and the heterogeneity of study populations and exercise programmes. This might explain why, for example, Pearsall et al. (2014) reported mixed findings from a recent systematic review and meta-analysis on RCTs, examining exercise interventions in those with serious mental health problems. This notwithstanding, there is broadly supportive evidence that exercise has therapeutic value for a range of mental health problems.

Additional benefits of physical activity in clinical mental health populations

The rationale for increasing physical activity within this population group extends beyond the psychological benefits. First, there are the physical health benefits. We have already noted the role of active lifestyles in promoting physical health and how those experiencing mental health problems are more likely to have comorbidities, such as diabetes and cardiovascular disease (Brown et al. 1999; Sokal et al. 2004). This is often the result of lifestyle behaviours, such as smoking, poor diet, and lack of exercise (Lambert et al. 2003; Wirshing 2004), and possible medication side-effects (Goff et al. 2005), which, again, affect this population group disproportionately. Second, and related to this, there is evidence that participation in physical activity, such as group exercise sessions, can improve motivation in mental health service users. Participants of these groups have reported feeling stimulated to take on other activities, even having increased impetus in other areas of their therapy (Fogarty and Happell 2005). This is a considerable benefit, given the common barriers of tiredness (from symptoms of disease or medication) and poor motivation, which prevent many from leading more healthy and active lives (Corry et al. 2004). Third, mental health problems are often associated with social exclusion and isolation (Hacking and Bates 2008). Again, evidence from studies of patients using adult mental health services has shown that physical activity groups can help to address unmet needs for company, daytime activities, and intimate relationships (Cleary et al. 2006). Case studies and interviews with service users have reported that physical activity groups can engage individuals with mental health services, promote normalising activities, and offer safe environments for social interaction (Carter-Morris and Faulkner 2003; Crone and Guy 2008; Faulkner and Sparkes 1999; Mental Health Foundation 2000; Wallcraft 1998). This can help to mitigate the risk of social isolation and stigma that is associated with impaired mental health.

For all of these reasons, it is not surprising that physical activity appears popular among people experiencing mental health problems. Over 75 per cent of those surveyed by the Social Exclusion Unit (2004) expressed a wish for better access to leisure services, which has been advocated by others who highlight the rights of service users to engage in such widely enjoyed recreational and leisure activities (Richardson et al. 2005).

Physical activity recommendations in mental health populations

Given the impact of inactivity on health, and the associated cost, HM Government (2014) recognises the importance of getting people active. Adults in the general population are recommended to accumulate at least 150 minutes of moderate-to-vigorous intensity activity weekly, ideally in bouts of ten minutes or more on a daily basis (Department of Health, 2011). There is also recognition that some groups in society are more likely to be inactive and therefore require specific, targeted approaches; this includes those with mental health problems. However, mental health problems often present in different ways, with symptoms of varying severity, which can differentially affect people's ability to engage in physical activity. As becomes clear later in this chapter, and as is highlighted in Chapter 11, each case must be treated individually, with a tailored approach specific to the individual. That said, earlier, we described how people with mental health problems can reap the same type of physical and mental health benefits from being physically active as the general population and, in fact, have greater needs given the increased presence of comorbidities. So it is logical that the current physical activity recommendations should be applicable both to those with and those without mental health problems (Rethorst and Trivedi 2013); the approach might be very different, but the target behaviour in terms of activity volume and intensity should be the same.

There is much debate around the type, intensity, and duration of physical activity to confer benefit in people with mental health problems. Findings from research in this area are by no means conclusive, limiting the evidence base to inform prescription guidelines (Morgan et al. 2013). To date, research has tended to focus on aerobic exercise, with few studies comparing, for example, aerobic, resistance, or mindful activity. Those that have report little difference (Rehor et al. 2001; Stathopoulou et al. 2006). Those undertaking an activity that is new to them (regardless of activity type) might benefit less because they need to learn a new skill (Rehor et al. 2001). A combination of aerobic- and resistance-based exercise has been proposed as preferable (Rethorst and Trivedi 2013) and feasible within a multidisciplinary team (Marzolini et al. 2009). However, it is noted that the type of exercise, and whether it is group or individual, might be less important than ensuring it is structured around their personal preferences (Morgan et al. 2013).

Research has focused on regular exercise, which is the ultimate behavioural target. However, within this population group, engagement in regular activity is not always feasible as opportunities for exercise may be limited, particularly for in-patients (Cormac et al. 2004). To better understand the effect of exercise (and the duration of benefits) some researchers have considered the impact of a single bout of activity. For some, where routine might be difficult, understanding the impact of just one session could prove useful in determining the frequency of sessions required to sustain benefit. Benefits have been shown immediately post exercise across a range of mental health diagnoses (Bartholomew et al. 2005; Ellis et al. 2013; Randall et al. 2014; Weinstein et al. 2010), with evidence that this might last for up to forty eight hours post-exercise (Randall, 2015). However, this was not supported elsewhere (Weinstein et al. 2010). Further research is required in this area to better support the structuring of exercise intervention.

Most evidence examines the impact of moderate-intensity exercise. This is a likely consequence of practicalities surrounding getting those experiencing mental health problems active, and the relatively low baseline activity levels from which many are starting. A minimum recommendation of thirty minutes of moderate-to-vigorous exercise three times per week has been suggested (Morgan et al. 2013). Although the authors advocate greater activity levels if possible, increasing to this level from a largely sedentary lifestyle will confer health benefits regardless of exercise intensity (Asztalos et al. 2010; Sparling et al. 2015; Wen et al. 2011). When considering exercise intensity, current fitness levels, and any comorbidity need to be taken into account, as exercise tolerance might be also lower in some mental health groups (Shah et al. 2007).

Physical activity in mental health services

The range of psychological, physical, and social benefits of physical activity make it an appealing and potentially low-cost mental health therapy (Daley 2002) that is often favourable to, or can provide a way to reduce, medication (Lautenschlager et al. 2004). The need for physical activity to be integrated into mental health services is recognised (Richardson et al. 2005; Walsh 2011). Health care professionals should be involved in the design and implementation of relevant activity programmes, and should include service users in the process (Cormac et al. 2004; Happell et al. 2012; NICE 2011, 2012). The Mental Health Foundation have recommended that all health professionals should understand the physical and psychological benefits that physical activity can offer, and should support people with mental health problems to become more active (Mental Health Foundation 2013b). Indeed, mental health nurses are often involved with intervention delivery, and are well placed to support patients with behaviour change interventions (Happell et al. 2012). These points are echoed by WHO (2013), who also recommend support from families and carers where appropriate.

Other important considerations for physical activity programmes within mental health services include, first, the need for regular communication between the exercise specialists and the care team (Cormac et al. 2004). Second, staff enthusiasm is key for successful engagement with, and encouragement of, service users (Cormac et al. 2004). Third, support networks are thought to play a fundamental role in any lifestyle behaviour change, particularly in this population group (NICE 2011). Where possible, these should be tailored to the individual, for example, the use of mobile devices and social media in young people (Killackey et al. 2011). Finally, basic practical considerations such as ensuring that people have appropriate clothes to exercise in is essential as they may have been an inpatient and sedentary for some time. A pre-exercise group is also suggested as it can provide patients with some of the skills and, therefore, increase confidence to exercise (Cormac et al. 2004).

Good practice for physical activity intervention in NHS mental health services

This section presents an example of good practice in an NHS mental health service. A number of stages are involved, from establishing an individual's suitability for physical activity intervention, to gathering of initial information, identifying appropriate physical activity, and case formation. This is designed to be a practical guide and offers some key points for practitioners to consider at each stage. Figure 5.1 shows an example care plan pathway to be followed by the exercise specialist, who should be trained as a level 4 specialist exercise instructor, specifically in Mental Health, on the Register of Exercise Professionals (REPs).

All individuals accessing NHS mental health services should receive a physical health screen prior to any formal physical activity engagement, to identify comorbidities and assess the initial disease risk classification of service users. This screening now follows the recommendations of the 2014 physical health CQUIN, (Commissioning for quality and innovation; NHS England/Commissioning Policy and Primary Care/Commissioning Policy and Resources, 2014; Shiers et al. 2014). Central to these recommendations is the need to develop, in partnership with the service user, a personalised evidence-based recovery plan that is well supported by a multi-disciplinary team (NHS England et al. 2014).

Triage information gathering

The exercise specialist determines if the referral is appropriate for physical activity therapy or whether a different discipline (physiotherapy, occupational therapy) is more appropriate, although a multi-disciplinary approach is likely. The exercise specialist should classify the priority of need (or risk level) for the identified problem and arrange to see the patient within a specific time period (Figure 5.1). The exercise specialist should make every effort to gather all pre-assessment information from the patient's GP, medical notes, family and friends, and significant others. Those closest to the individual will often have a valuable understanding of the patient and information

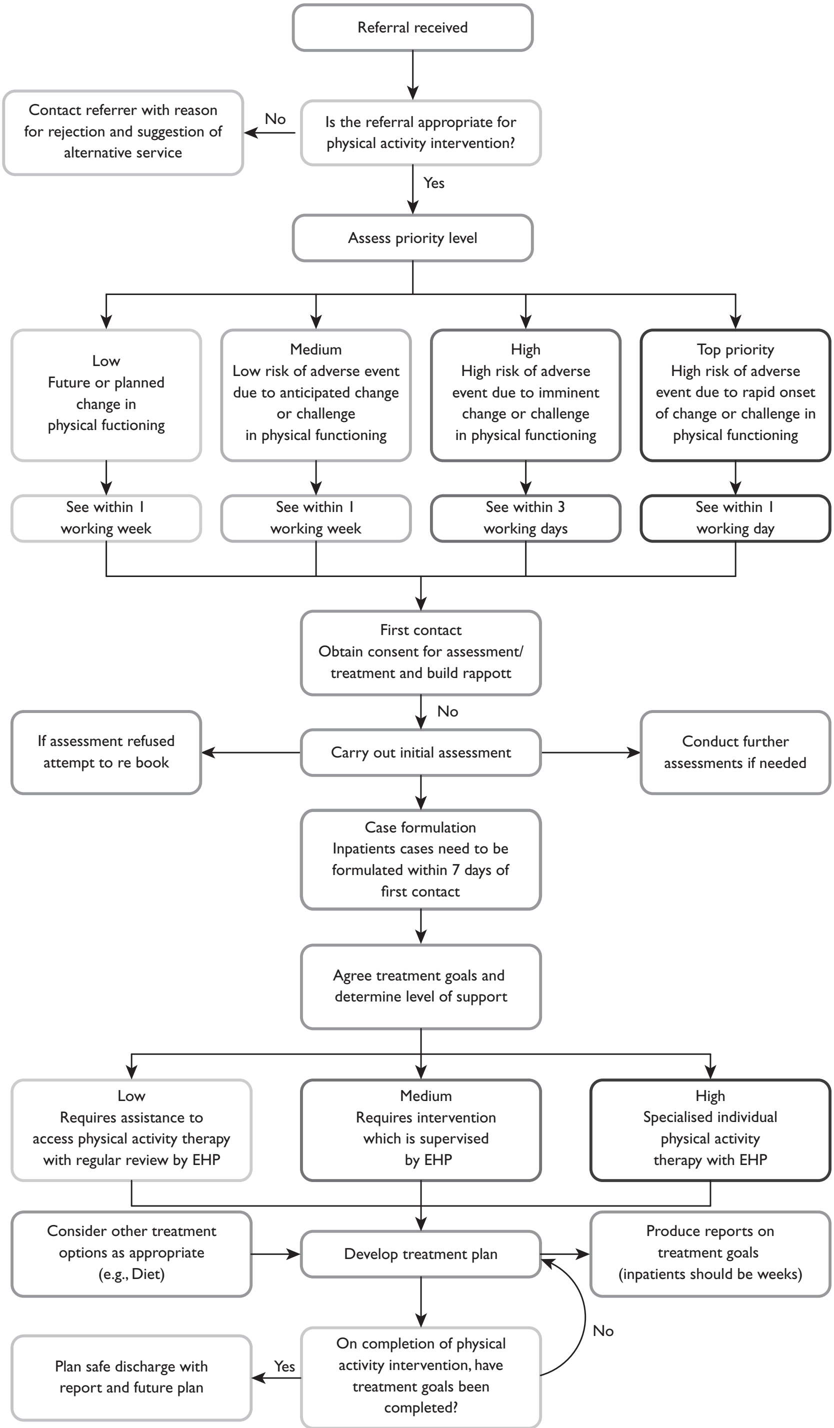


Figure 5.1 Referral pathway.

not otherwise disclosed to health care professionals. It is anticipated that the exercise specialist will have a thorough understanding of the concerns and needs of the patient prior to progressing to the next stage, 'first contact'.

First contact

At first contact, the exercise specialist should build a rapport with the patient, aiming to understand their interests and desire to engage in the process. It may also include further discussions with carers, relatives, and relevant professions to follow-up on any concerns arising at triage. At this point, patient consent is sought for subsequent assessment and treatment. The patient may not wish to engage, but this should not preclude further involvement. It is recognised that people with mental health problems are likely to have reduced levels of motivation and the exercise specialist should periodically re-engage with them as appropriate.

Assessment

Within one-week of first contact, there should be an initial assessment, completed in collaboration with the patient. This might take place over a number of meetings, dependent on the patient and the extent of their needs. The aim is to generate sufficient information for case formulation, so it requires consideration of both psychiatric and physical health concerns. The Lester Screening Tool offers a useful way to identify the physical health concerns of people with psychiatric illnesses, and presents guidance for appropriate interventions (Shiers et al. 2014). There is the potential to prioritise physical health needs at this stage. But, because the patient's mental illness will determine their ability to tackle such problems, practitioners should give sufficient credence to the mental health facets. This requires that practitioners understand patients' clinical diagnoses and identify constraints to participation in any physical activity intervention.

Case formulation

The exercise specialist should consider the level of support needed to address both the physical and mental health concerns, which are determined by the patient's classified risk (from triage and subsequent information gathering). At this stage, the exercise specialist should consider the design of an exercise and health care plan to guide the intervention. Exercise-specialist care plans are an integral component to the safe delivery of physical activity intervention and should reflect the specific needs of the individual. Principles of re-motivation and motivational interviewing, widely-used practices in the mental health arena (described in Chapter 13), might prove useful at this stage.

Depending on risk level and patient motivation, the intervention might take a number of forms. It could involve a supervised exercise specialist-led intervention, which involves a one-to-one specific structured intervention, including a thorough assessment of physiological and psychological changes, whereby the exercise specialist continually tailors the activity to the individual's needs (Figure 5.1; Box 5.1). However, such specialised supervised intervention might not be appropriate or necessary in an in-patient psychiatric hospital setting and, for many, physical activity interventions can be conducted within their own local community.

Box 5.1 Case study

The following case study describes an individual who has received an exercise and health intervention within an NHS mental health service in the UK.

The individual. Bobby is a twenty three year-old male who has been detained under Section 2 of the Mental Health Act (1983). This was his first admission to hospital and he was previously unknown to the local mental health service. Bobby was experiencing thoughts of paranoia, and was suspicious and guarded of those around him. He was continually seeking to leave the hospital. He has been referred to the exercise specialist for an assessment of his physical health needs as standard procedure following admission and he has indicated an interest in physical activity.

Triage gathering. His brother reported that Bobby had no known physical health concerns, but had not been screened or seen his GP for over three years. He stated that Bobby was a keen athlete until recently and regularly attended a local athletics club. Recently, Bobby's odd behaviour caused friends to distance themselves, leaving him isolated within the family home. The GP confirmed that there were no known physical health or heredity concerns.

First contact. First contact was made in the hospital ward. Bobby was obviously disturbed in his thought processes as he was observed to be responding to some internal stimuli. He was, however, able to respond to verbal direction from the exercise specialist and, whilst suspicious at first, appeared reassured and interested when given an explanation of exercise specialist role. Bobby confirmed the information provided by his brother, giving details of his athletics and physical activity participation. Bobby reported that he was predominantly interested in his physique and appearance, and regularly attended his local gym for resistance training. Bobby expressed motivation to participate in any physical activity intervention but wanted this to occur within his local community rather than the hospital.

Exercise health practitioner assessment. Bobby was happy to cooperate with all aspects of the assessment; however, given the disturbance in his thought processes, it was not practical to complete in one session. Consideration was given to Bobby's level of understanding at this time. Several visits also provided the opportunity to build rapport between practitioner and patient, ensuring Bobby's needs were fully considered.

Metabolic considerations
Smoking: < 10cigarettes a day
Lifestyle: happy with his own diet (predominantly protein-orientated)
Body Mass Index: 25–30 $kg \cdot m^2$
Blood pressure: < 130/90 mmHg
Glucose regulation: < 5 $mmol \cdot L^{-1}$
Blood lipids: total cholesterol < 5 $mmol \cdot L^{-1}$

Mental health consideration. Whilst there was no formal diagnosis, Bobby's symptoms were similar to those of a psychotic-type illness (e.g., hallucinations, delusions, confused and disturbed thoughts, lack of insight and self-awareness). Further mental health consideration was Bobby's continual ambition to leave hospital, yet legally detained under Section 2 of the Mental Health Act. This means he was detained for up to twenty eight days for a mental health assessment and any treatment required.

Case formulation. Bobby's care plan was completed. All information was kept on record, reviewed, and updated to ensure the care plan remained relevant to his needs.

Development. Bobby engaged in the physical activity intervention on a 1:1 basis for two weeks with a focus on weight training. He remained conscious of his appearance and how others perceived him. Whilst it was assumed this was attributable to his mental illness, it became evident that Bobby had always held these thoughts. During the course of intervention, Bobby expressed additional interests, specifically in swimming. Although still detained under Section 2, Bobby had been prescribed Section 17 leave under the Mental Health Act, which allowed periods of supervised leave outside the hospital. Whilst not visibly at ease within a group environment, it was deemed fitting to offer Bobby the opportunity to participate in the swimming group in the local community, as he had not absconded from any previous leave and had formed good relations with the exercise specialist. Furthermore, he was now accepting his psychotropic medication. Consequently, Bobby used several periods of escorted leave to attend the local swimming facility. Over time, he was observed to be more trusting of others and able to interact socially within a group setting. Thereafter, Bobby was able to attend both individual and group intervention without the need for direct exercise specialist supervision or mental health escort. Bobby was subsequently discharged from hospital and referred to his local community mental health team.

Progression. Following discharge, Bobby did not seek the support services of his community mental health team, specifically his community psychiatric nurse. However, he remained keen to engage with available physical activity interventions and was referred to the exercise specialist attached to the local community mental health team. At this time, Bobby was avoiding any association with his mental health concerns. It was, therefore, important to extend the exercise specialist relationship to safeguard against deterioration in his mental health problem. Although exercise specialist intervention continued, Bobby was disengaged from other services and in denial of his psychiatric disorder. Consequently, his mental health began to deteriorate and he was subsequently admitted to hospital under Section 3 of the Mental Health Act. Someone detained under section 3 is done so if they are known to mental health services and require treatment in hospital for their own health or for the protection of other people.

Review. Bobby had become well known to his mental health services and occasionally needed admission to hospital. Whilst his mental health disorder impeded his thought processes, his previous relationships with exercise specialists ensured that he accepted the services provided. The exercise specialist continued to play a significant role in Bobby's acceptance of his mental health disorder.

For those not considered high risk, but who require supervised physical activity within the institutional setting, there are some important considerations. Some psychiatric hospitals fail to address the need for physical activity provision. Authors have suggested that management style, ethos, and culture in such places can contribute to socially-disengaged and inactive in-patients (Radcliffe and Smith 2007). Where barriers exist, one solution is that the activities are prescribed by the exercise specialist, but delivery is facilitated by other practitioners. This can afford greater frequency and flexibility of programmes, but requires that exercise specialists involve other professions in the design and creation of physical activity interventions.

Finally, those with lower risk levels and sufficient motivation might be able to implement their own activity plan, with regular reviews from the exercise specialist. In these instances, activity is most likely to occur within the community. While there are benefits, potential barriers to community-based physical activity to consider include social exclusion, access to facilities, and stigma (Social Exclusion Unit 2004). Support mechanisms to address such issues might involve buddying, a practice commonly used by mental health providers, which offers both emotional and practical peer support to help with engagement (e.g. introducing individuals to new groups or facilities, offering encouragement to attend sessions).

In all cases, ongoing assessment and review are important to ensure activities are appropriate and tailored to individual need, ultimately with a view to safe discharge with a plan for continued activity.

Future research

Several areas warrant further research:

- Greater understanding of the duration of psychological benefits of exercise would provide practitioners with the evidence base to inform the frequency of sessions;
- Use of objective physical activity measurement (e.g. accelerometry) would provide researchers with a more valid measure of physical activity frequency, intensity, and duration to allow exploration of dose-response in different conditions. However, it is acknowledged that using activity monitors is likely to present challenges beyond those experienced in studies of non-clinical populations, given the nature of certain mental health conditions (e.g. patients experiencing symptoms of paranoia);
- The challenges of working with populations who have clinical mental health diagnoses often result in studies and evaluations with relatively small sample sizes and short follow-up periods. Efforts to address this common limitation might require closer collaboration between mental health service providers and academic institutions to allow more routine and longer-term data collection.

Chapter summary

This chapter has introduced the scale of the challenge that mental health problems pose, globally and in the UK, in terms of health, health care provision, and the economic burden. The multiple benefits of physical activity for mental, physical, and social health, and the absence of side-effects, make it an appealing adjunct or alternative to pharmacological treatments. Although there are some limitations and inconsistencies in the evidence base, there is sufficient support to implement exercise interventions in those with mental health problems. We have tried to emphasise that the range and complexities of mental illnesses demand each individual be assessed and treated in a way that is appropriate to their need. As discussed in other chapters of this book, effecting behaviour change in the general population is a considerable challenge. With the additional complications and challenges that mental health problems present, effective intervention in this population group warrants careful consideration. The good practice example and case study outline some practical steps for practitioners and those involved in the design and delivery of mental health services, to help with the basic processes and factors for including physical activity as part of the care pathway in mental health service users.

Further reading

Cormac, I., Martin, D. and Ferriter, M. (2004). Improving the physical health of long-stay psychiatric in-patients. *Advances in Psychiatric treatment*, 10, 107–115.

Mental Health Foundation (2013) *Starting today: the future of mental health services*. London, Mental Health Foundation.

Pearsall, R., Smith, D. J., Pelosi, A. and Geddes, J. (2014). Exercise therapy in adults with serious mental illness: a systematic review and meta-analysis. *BMC Psychiatry*, 14, 117, doi:10.1186/1471-244X-14-117.

World Health Organisation (2013). *Mental health action plan 2013–2020*. Geneva, World Health Organisation.

Study tasks

1 Read the most recent mental health guidance documents (e.g. by organisations such as NICE, Mind, WHO, Mental Health Foundation) and summarise their advice relating to physical activity and exercise.

2 Choose a type of mental health problem (e.g. depression, schizophrenia) and review the evidence, commenting on both the strength of support for exercise interventions and a key area for future research.

3 Within your local area, identify what physical activity opportunities are available for mental health service users, and consider what challenges should be considered if trying to encourage attendance.

References

Allender, S., Foster, C., Scarborough, P. and Rayner, M. (2007). The burden of physical activity-related ill health in the UK. *Journal of Epidemiology and Community Health*, 61, 344–8.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Ansseau, M., Dierick, M., Buntinkx, F., Cnockaert, P., De Smedt, J., Van Den Haute, M. and Vander Mijnsbrugge, D. (2004). High prevalence of mental disorders in primary care. *Journal of Affective Disorders*, 78(1), 49–55.

Asztalos, M., De Bourdeaudhuij, I. and Cardon, G. (2010). The relationship between physical activity and mental health varies across activity intensity levels and dimensions of mental health among women and men. *Public Health Nutrition*, 13(8), 1207–14.

Bartholomew, J. B., Morrison, D. and Ciccolo, J. T. (2005). Effects of acute exercise on mood and well-being in patients with major depressive disorder. *Medicine and Science in Sports and Exercise*, 37(12), 2032–7.

Biddle, S. J. H. and Mutrie, N. (2008). *Psychology of physical activity: determinants, well-being and interventions*. Abingdon: Routledge.

British Heart Foundation National Centre (2013). *Economic costs of physical inactivity*. Loughborough: British Heart Foundation National Centre.

Brown, H. E., Pearson, N., Braithwaite, R. E., Brown, W. J. and Biddle, S. J. H. (2013). Physical Activity Interventions and Depression in Children and Adolescents. *Sports Medicine*, 43(3), 195–206.

Brown, S., Barraclough, B. and Inskip, H. (2000). Causes of the excess mortality of schizophrenia. *British Journal of Psychiatry*, 177, 212–17.

Brown, S., Birtwistle, J., Roe, J. and Thompson, C. (1999). The unhealthy lifestyle of people with schizophrenia. *Psychological Medicine*, 29, 697–701.

Brown, W. J., Ford, J. H., Burton, N. W., Marshall, A. L. and Dobson, A. J. (2005). Prospective study of physical activity and depressive symptoms in middle-aged women. *American Journal of Preventive Medicine*, 29(4), 265–72.

Camacho, T. C., Roberts, R. E., Lazarus, N. B., Kaplan, G. A. and Cohen, R. D. (1991). Physical activity and depression: evidence from the Alameda County Study. *American Journal of Epidemiology*, 134(2), 220–31.

Carless, D. and Douglas, K. (2004). A golf programme for people with severe and enduring mental health problems. *Journal of Health Promotion*, 3(4), 26–39.

Carter-Morris, P. and Faulkner, G. (2003). A football project for services users: the role of football in reducing social exclusion. *Journal of Mental Health Promotion*, 2(2), 24–30.

Cassidy, K., Kotynia-English, R., Acres, J., Flicker, L., Lautenschlager, N. T. and Almeide, O. P. (2004). Association between lifestyle factors and mental health measures among community dwelling older women. *Australian and New Zealand Journal of Psychiatry*, 38, 940–47.

Centre for Mental Health (2011). *The economic and social costs of mental health problems in 2009/10*. Accessed via: www.centreformentalhealth.org.uk/pdfs/economic_and_social_costs_2010.pdf.

Chen, J. and Millar, W. J. (1999). Health effects of physical activity. *Health Reports*, 11(1), 21–30.

Chue, P. and Kovacs, C. S. (2003). Safety and tolerability of atypical antipsychotics in patients with bipolar disorder: prevalence, monitoring and management. *Bipolar Disorders*, 5 (suppl. 2), 62–79.

Cleary, M., Freeman, A., Hunt, G. E. and Walter, G. (2006). Patient and carer perceptions of need and associations with care-giving burden in an integrated adult mental health service. *Social Psychiatry and Psychiatric Epidemiology*, 41(3), 208–14.

Connolly, M. and Kelly, C. (2005). Lifestyle and physical health in schizophrenia. *Advances in Psychiatric Treatment*, 11, 125–32.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., McMurdo, M. and Mead, G. E. (2013). Exercise for depression. *Cochrane Database Systematic Review*, 9:CD004366.

Cormac, I., Martin, D. and Ferriter, M. (2004) Improving the physical health of long-stay psychiatric in-patients. *Advances in Psychiatric treatment*, 10, 107–15.

Corry, P., Dru Drury, C. and Pinfold, V. (2004). *Lost and found. Voices from the forgotten generation*. London: Rethink.

Craft, L. L. and Landers, D. M. (1998). The effect of exercise on clinical depression and depression resulting from mental illness: a meta-analysis. *Journal of Sport and Exercise Psychology*, 20, 339–57.

Craft, L. L. and Perna, F. M. (2004). The benefits of exercise for the clinically depressed. *Primary Care Companion Journal of Clinical Psychiatry*, 6(3), 104–13.

Cramer, J. A. and Rosenbeck, R. (1998). Compliance with medication regimes for mental and physical disorders. *Psychiatric Services*, 49(2), 196–201.

Crone, D. and Guy, H. (2008). 'I know it's only exercise, but to me it is something that keeps me going': a qualitative approach to understanding mental health service users experiences of sports therapy. *International Journal of Mental Health Nursing*, 17(3), 197–207.

Daley, A. (2002) Exercise therapy and mental health in clinical populations: is exercise therapy a worthwhile intervention? *Advances in Psychiatric Treatment*, 8(4), 262–70.

Davies, S. C. (2014). *Annual report of the Chief Medical Officer, surveillance volume 2012: on the state of the public's health*. London: Department of Health.

Department of Health. (2001). *Making it happen: a guide to delivering mental health promotion*. London: HMSO.

Department of Health (2004). *At least five a week*. London: HM Government.

Department of Health (2011). *Start active, stay active: a report on physical activity for health from the four home countries' Chief Medical Officers*. London: HM Government.

Dimeo, F., Bauer, M., Varahram, I., Proest, G. and Halter, U. (2001). Benefits from aerobic exercise in patients with major depression: a pilot study. *British Journal of Sports Medicine*, 35, 114–17.

Ehmann, T. and Hanson, L. (2002). *Early psychosis: a care guide summary*. Vancouver: The University of British Columbia.

Ellis, N. J., Crone, D., Davey, R. and Grogan, S. (2007). Exercise interventions as an adjunct therapy for psychosis: A critical review. *British Journal of Clinical Psychology*, 46(1), 95–111.

Ellis, N. J., Randall, J. A. and Punnett, G. (2013). The effects of a single bout of exercise on mood and self-esteem in clinically diagnosed mental health patients. *Open Journal of Medical Psychology*, 2(3), 81–5.

Faulkner, G. and Sparkes, A. (1999). Exercise as a therapy for schizophrenia: an ethnographic study. *Journal of Sport and Exercise Psychology*, 21, 52–69.

Faulkner, G. and Taylor, A. H. (2005). Exercise and mental health promotion. In G. E. J. Faulkner and A. H. Taylor, eds., *Exercise, health and mental health: emerging relationships*. London: Routledge. Ch. 1.

Fogarty, M. and Happell, B. (2005). Exploring the benefits of an exercise program for people with schizophrenia: a qualitative study. *Issues in Mental Health Nursing*, 26, 341–51.

Forbes, D., Thiessen, E. J., Blake, C. M., Forbes, S. C. and Forbes, S. (2013). Exercise programs for people with dementia. *The Cochrane Library*, 12:CD006489.

Goff, D. C., Cather, C., Evins, E., Henderson, D. C., Freudenreich, O., Copeland, P. M., Bierer, M., Duckworth, K. and Sacks, F. M. (2005). Medical morbidity and mortality in schizophrenia: guidelines for psychiatrists. *Journal of Clinical Psychiatry*, 66(2), 183–94.

Hacking, S. and Bates, P. (2008). The inclusion web as a tool for person-centred planning and service evaluation. *Mental Health Review Journal: Research, Policy and Practice*, 13, 4–15.

Haddad, P. (2005). Weight change with atypical antipsychotics in the treatment of schizophrenia. *Journal of Psychoparmacology*, 19(6), 16–27.

Happell, B., Davies, C. and Scott, D. (2012). Health behaviour interventions to improve physical health in individuals diagnosed with a mental illness: a systematic review. *International Journal of Mental Health Nursing*, 21, 236–47.

Harris, E., C. and Barraclough, B. (1998). Excess mortality of mental disorder. *The British Journal of Psychiatry*, 173, 11–53.

Harvey, S. B., Hotopf, M., Øverland, S. and Mykletun, A. (2010). Physical activity and common mental disorders. *The British Journal of Psychiatry*, 197, 357–64.

Haslam, C., Brown, S., Atkinson, S. and Haslam, R. (2004). Patients' experiences of medication for anxiety and depression: effects on working life. *Family Practice*, 21(2), 204–12.

Hassmen, P., Koivula, N. and Uutela, A. (2000). Physical exercise and psychological well-being: a population study in Finland. *Preventive Medicine*, 30, 17–25.

Health and Social Care Information Centre (2012). *Health Survey for England 2012: health, social care and lifestyles*. London: HSCIC.
Heyn, P., Abreu, B. C. and Ottenbacher, K. J. (2004). The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Archives of Physical Medicine and Rehabilitation*, 85(10), 1694–1704.
HM Government (2014). *Moving more, living more: the physical activity Olympic and Paralympic legacy for the nation*. London: HM Government.
Holley, J., Crone, D., Tyson, P. and Lovell, G. (2011). The effects of physical activity on psychological well-being for those with schizophrenia: a systematic review. *British Journal of clinical Psychology*, 50(1), 84–105.
Hutchinson, D. S. (2005). Structured exercise for persons with serious psychiatric disabilities. *Psychiatric Services*, 56(3), 353–54.
Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R. and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*. 62(6), 593–602.
Killackey, E., Anda, A. L., Gibbs, M., Alvarez-Jimenez, M., Thompson, A., Sun, P. and Baksheev, G.N. (2011). Using internet enabled mobile devices and social networking technologies to promote exercise as an intervention for young first episode psychosis patients. *BMC Psychiatry*, 11, 80 doi:10.1186/1471-244X-11-80
Lam, R. and Kennedy, S. H. (2004). Evidence-based strategies for achieving and sustaining full remission in depression: focus on metaanalyses. *Canadian Journal of Psychiatry*, 49(suppl.1), 17–26S.
Lambert, T. J. R., Velakoulis, D. and Pantelis, C. (2003). Medical comorbidity in schizophrenia. *Medical Journal of Australia*, 178(suppl.), S67–70.
Lampinen, P., Heikkinen, R. L. and Ruoppila, I. (2000). Changes in intensity of physical exercise as predictors of depressive symptoms among older adults: an eight-year follow-up. *Preventive Medicine*, 30, 371–80.
Lautenschlager, N. T., Almeido, O. P., Klicker, L. and Janca, A. (2004). Can physical activity improve the mental health of older adults? *Annals of General Hospital Psychiatry*, 3(12), doi: 10.1186/1475-2832-3-12.
Law, L. L. F., Barnett, F., Yau, M. K. and Gray, M. A. (2014). Effects of combined cognitive and exercise interventions on cognition in older adults with and without cognitive impairment: a systematic review. *Ageing Research Reviews*, 15, 61–75.
Malchow, B., Reich-Erkelenz, D., Oertel-Knöchel, V., Keller, K., Hasan, A., Schmitt, A., Scheewe, T. W., Cahn, W., Kahn, R. S. and Falkai, P. (2013). The effects of physical exercise in schizophrenia and affective disorders. *European Archives of Psychiatry* and *Clinical Neuroscience*, 263(6), 451–67.
Mammen, G. and Faulkner, G. (2013). Physical activity and the prevention of depression: a systematic review of prospective studies. *American Journal of Preventive Medicine*, 45(5), 649–57.
Martinsen, E. W. (2000) Physical activity for mental health. *Tidaakr nor Laegeforen*, 120(25), 3054–6.
Marzolini, S., Jensen, B. and Melville, P. (2009). Feasibility and effects of a group-based resistance and aerobic exercise program for individuals with severe schizophrenia: a multidisciplinary approach. *Mental Health and Physical Activity*, 2(1), 29–36.
McCrone, P., Dhanasiri, S., Patel, A., Knapp, M. and Lawton-Smith, S. (2008). *Paying the price. The cost of mental health care in England to 2026*. London: The King's Fund.
McDevitt, J., Wilbur, J., Kogan, J. and Briller, J. (2005). A walking program for outpatients in psychiatric rehabilitation: a pilot study. *Biological Research for Nursing*, 7(2), 87–97.
Mead, G. E., Morley, W., Campbell, P., Greig, C. A., McMurdo, M. and Lawlor, D. A. (2008) Exercise for depression. *Cochrane Database of Systematic Reviews*, 4, CD004366.

Mehlig, K., Skoog, I., Waern, M., Miao Jonasson, J., Lapidus, L., Björkelund, C., Ostling, S. and Lissner, L. (2014). Physical activity, weight status, diabetes and dementia: a 34-year follow-up of the Population Study of Women in Gothenburg. *Neuroepidemiology*, 42(4), 252–9.

Mental Health Foundation. (2000). *Strategies for living: a report of user-led research into people's strategies for living with mental distress*. London: Mental Health Foundation.

Mental Health Foundation (2013a). *Starting today: the future of mental health services*. London, Mental Health Foundation.

Mental Health Foundation (2013b). *Let's get physical: the impact of physical activity on wellbeing*. London, Mental Health Foundation.

Mentality (2004). *Mental health promotion – implementing standard one of the National Service Framework for Mental Health*. London: SCMH.

Mind. (2007). *Ecotherapy: the green agenda for mental health*. London: Mind.

Morgan, A. J., Parker, A. G., Alvarez-Jimenez, M. and Jorm, A. F. (2013). Exercise and mental health: an Exercise and Sports Science Australia commissioned review. *Journal of Exercise Physiology online*, 16(4), 64–73.

National Institute for Clinical Excellence (NICE) (2011). *Service user experience in adult mental health: improving the experience of care for people using adult NHS mental health services*. NICE guidelines [CG136].

National Institute for Clinical Excellence (NICE) (2012). *Quality standard for patient experience in adult NHS services*. NICE guidelines [QS15].

NHS England/Commissioning Policy and Primary Care/Commissioning Policy and Resources (2014), *Commissioning for Quality and Innovation (CQUIN): 2014/15 guidance*. Leeds: Commissioning Policy and Primary Care.

NHS Information Centre for Health and Social Care. (2009). *Health survey for England 2008: physical activity and fitness*. London: NHS Information Centre for Health and Social Care.

Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K. and Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *The Lancet Neurology*, 13(8), 788–94.

O'Kane, P. and McKenna, B. (2002). Five-a-side makes the difference. *Mental Health Nursing*, 22(5), 6–9.

Ossa, D. and Hutton, J. (2002). *The economic burden of physical inactivity in England*. London: MEDTAP International.

Pasco, J. A., Williams, L. J., Jacka, F. N., Henry, M. J., Coulson, C. E., Brennan, S. L., Leslie, E., Nicholson, G. C., Kotowicz, M. A. and Berk, M. (2011). Habitual physical activity and the risk for depressive and anxiety disorders among older men and women. *International Psychogeriatrics*, 23(2), 292–8.

Pearsall, R., Smith, D. J., Pelosi, A. and Geddes, J. (2014). Exercise therapy in adults with serious mental illness: a systematic review and meta-analysis. *BMC Psychiatry* 14, 117, doi: 10.1186/1471-244X-14-117.

Pelham, T. W. and Campagna, P. D. (1991). Benefits of exercise in psychiatric rehabilitation of persons with schizophrenia. *Canadian Journal of Rehabilitation*, 4(3), 159–68.

Pelham, T. W., Campagna, P. D., Ritvo, P. G. and Birnie, W. A. (1993). The effects of exercise therapy on clients in a psychotic rehabilitation program. *Psychosocial Rehabilitation Journal*, 16, 75–84.

Pilgrim, D. (2005). *Key concepts in mental health*. London: SAGE.

Priest, P. (2007). The healing balm effect; using a walking group to feel better. *Journal of Health Psychology*, 12(1), 36–52.

Radcliffe, J. and Smith, R. (2007). Acute in-patient psychiatry: how patients spend their time on acute psychiatric wards. *Psychiatric Bulletin*, 31, 167–70.

Randall, J. (2015). *The effect of a single bout of exercise on mood and self-esteem in individuals with a clinical mental health diagnosis*, Doctoral research, Staffordshire University.

Randall, J., Ellis, N., Gidlow, C. and Jones, M. (2014). Comparing mental health diagnoses: changes in mood and self-esteem following a single bout of exercise. *The Journal of Psychological Therapies in Primary Care*, 3(1), 34–6.

Rehor, P. R., Dunnagan, T., Stewart, G. and Cooley, D. (2001). Alteration of mood state after a single bout of non-competitive and competitive exercise programs. *Perceptual and Motor Skills*, 93, 249–56.

Rethorst, C. D. and Trivedi, M. H. (2013). Evidence-based recommendations for the prescription of exercise for major depressive disorder. *Journal of Psychiatric Practice,* 19(3), 204–12.

Richardson, C. R., Faulkner, G., McDevitt, J., Skrinar, G. S., Hutchinson, D. S. and Piette, J. D. (2005). Integrating physical activity into mental health services for persons with serious mental illness. *Psychiatric Services*, 56(3), 324–31.

Royal College of Psychiatrists (2010). *No health without mental health: the supporting evidence*. Royal College of Psychiatrists, London.

Scarborough, P., Bhatnagar, P., Wickramsinghe, K. K., Allender, S., Foster, C. and Rayer, M. (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to the 2006–2007 NHS costs. *Journal of Public Health*, 33(4) 527–35.

Shah, A., Alshaher, M., Dawn, B., Siddiqui, T., Longaker, R., Stoddard, M. F. and El-Mallakh, R. (2007). Exercise tolerance is reduced in bipolar illness. *Journal of Affective Disorders*, 104, 191–5.

Shiers, D. E., Rafi, I., Cooper, S. J., Holt, R. I. G. (2014). *Positive cardiometabolic health resource: an intervention framework for patients with psychosis and schizophrenia. 2014 update*. London: Royal College of Psychiatrists.

Sin, J. and Gamble, C. (2003). Managing side-effects to the optimum: valuing a client's experience. *Journal of Psychiatric and Mental Health Nursing*, 10(2), 147–53.

Singleton, N., Bumpstead, R., O'Brien, M., Lee A. and Meltzer, H. (2001). *Psychiatric morbidity among adults living in private households, 2000*. London: The Stationary Office.

Social Exclusion Unit (2004). *Mental health and social exclusion*. London: Office of the Deputy Prime Minister.

Sokal, J., Messias, E., Dickerson, F. B., Kreyenbuhl, J., Brown, C. H., Goldberg, R. W. and Dixon, L. B. (2004). Co-morbidity of medical illnesses among adults with serious mental illness who are receiving community psychiatric services. *Journal of Nervous Mental Diseases*, 192(6), 421–7.

Sparling, P. B., Howard, B. J., Dunstan, D. W. and Owen, N. (2015). Recommendations for physical activity in older adults. *British Medical Journal*, 350, h100.

Stathopoulou G., Powers, M. B., Berry A. C., Smits, J. A. J. and Otto, M. W. (2006). Exercise interventions for mental health: a quantitative and qualitative review. *Clinical Psychology (New York)*, 13, 179–93.

Strawbridge, W. J., Deleger, S., Roberts, R. E. and Kaplan, G. A. (2002). Physical activity reduces the risk of subsequent depression for older adults. *American Journal of Epidemiology*, 156(4), 328–34.

Ten Have, M., de Graaf, R. and Monshouwer, K. (2011) Physical exercise in adults and mental health status: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Journal of Psychosomatic Research*, 71, 342–8.

The MaGPIe Research Group (2003). The nature and prevalence of psychological problems in New Zealand primary care: a report on Mental Health and General Practice Investigation (MaGPIe). *The New Zealand Medical Journal*, 116, 1171.

Vancampfort, D., Probst, M., Helvik Skjaerven, L., Catalán-Matamoros, D., Lundvik-Gyllensten, A., Gómez-Conesa, A., Ijntema, R. and De Hert, M. (2012). Systematic review of the benefits of physical therapy within a multidisciplinary care approach for people with schizophrenia. *Physical Therapy*, 92, 11–23.

Wallcraft, J. (1998). *Healing minds: a report on current research, policy and practice concerning the use of complementary and alternative therapies for a wide range of mental health problems*. London: Mental Health Foundation.

Walsh, R. (2011). Lifestyle and mental health. *American Psychologist*, 66(7), 579–92.

Weinstein, A. A., Deuster, P. A., Francis, P. L., Beadling, C. and Kop, W. J. (2010). The role of depression in short-term mood and fatigue responses to acute exercise. *International Journal of Behavioral Medicine*, 17(1), 51–7.

Wen, C. P., Wai J. P. M., Tsai, M. K., Yi Yang, C., Cheng, T. Y. D., Lee, M.-C., Chan, H. T., Tsao, C. K., Tsai, S. P. and Wu, X. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *The Lancet*, 378(9798), 1244–53.

Wirshing, D. A. (2004). Schizophrenia and obesity: impact of antipsychotic medications. *Journal of Clinical Psychiatry*, 65(18), 13–26.

Wolff E., Gaudlitz K., von Lindenberger B. L., Plag J., Heinz A., and Ströhle, A. (2011). Exercise and physical activity in mental disorders. *European Archives of Psychiatry and Clinical Neuroscience*, 261(suppl. 2), S186–91.

World Health Organisation (WHO) (2008). *Global burden of disease, 2004 update*. Geneva, World Health Organisation.

World Health Organisation (2010). *Global recommendations for physical activity for health*. Geneva, World Health Organisation.

World Health Organisation (2013). *Mental health action plan 2013–2020*. Geneva, World Health Organisation.

Chapter 6

The injury risk profiling process

David Joyce and Dan Lewindon

Introduction

For centuries, humans have sought to predict the future in order to ease anxiety about prospective events. Nowadays, this ambition extends to sport and exercise, where performance support teams employ a series of tests in the hope of gaining an insight into future athletic performance or injury risk.

Originally, the purposes of pre-participation evaluations in sport were to assure the coaching staff that their players had commenced the competitive season with a common level of health and fitness and to identify 'treatable' conditions. Whilst these objectives are still present, the scope and ambition of screening has extended to include:

- identification of areas of weakness or pain associated with performance to provide a platform for injury risk management programmes;
- assessment of physical capacities (e.g. aerobic power, acceleration, vertical jump height);
- determination of benchmarks (e.g. what is normal for a certain athletic population based upon age, playing position, etc.);
- assessment of recovery from previous injury (when it is used as a baseline measurement);
- elimination of intrinsic faults prior to coaching technique/training programme changes (such as back squat proficiency or plié technique).

The focus of this chapter is not so much to detail athletic performance testing, as this is covered elsewhere in the text, but rather to demonstrate how a true interdisciplinary approach to athletic profiling can help identify the athlete at risk of injury so that plans can be put in place to mitigate that danger.

The aim of assessing every single aspect of an athlete is unrealistic and not supported by scientific research. The field of tests that *could* be applied is too vast, many of which will be irrelevant to the athlete being tested. It is clear that a strategic approach is required, one that relies on the gathering of intelligence as a method of sifting through all the possible things that 'could go wrong', in order to hone in on those things that are 'most likely to'. In this way, it can be seen that the principal aim of the screening process is for it to yield meaningful *information*, rather than just data.

What is also clear is that screening should not be used as the definitive assessment tool *per se*, but as a method of filtering through all the possible injury scenarios to highlight certain areas of an athlete's profile that require management or further investigation.

This chapter will present a framework for the development of a pragmatic, evidence-based screening tool that is functionally relevant and individually targeted, the fundamentals of which can be applied across all sports and age groups, as well being useful in industrial and military environments.

Disrupting the injury equation

The occurrence of an injury has wide-ranging effects on the athlete, including:

- reducing time spent training or competing
- imposing a negative effect on sporting performance
- financial costs and psychological implications.

Indirect costs of injury are also evident with impacts sometimes being felt by their team, their family and the wider community. It is clear, therefore, that prevention of injuries in sport and exercise is a worthwhile aim. It is unrealistic to expect all injuries to be preventable. What we can aspire to, however, is a situation where the risk of injury is reduced.

The premise for using screening as a keystone of injury risk reduction relies upon it identifying individuals susceptible to injury. What turns an athlete into a susceptible one is best seen in the following equation:

Injury risk equation 1

Predisposed individual + extrinsic risk factors
= susceptible individual

A predisposed individual is one who is more exposed to injury due to their own intrinsic risk profile. Such intrinsic risks may include past injury, age, reduced joint range of motion (ROM) and muscle weakness. If the predisposed athlete is exposed to extrinsic risk factors, they then become susceptible to injury. Extrinsic risk factors are those that are applied to the athlete, such as training regimen, weather and the nature and laws of the sport itself.

An example of this is a female netballer with hamstring weakness who has poor landing mechanics. In this instance, the athlete is predisposed to knee injuries due to her gender and kinematics, both of which are linked with knee injury risk.[1] Her sport then exposes her further to risk due to the rules promoting sudden decelerations and pivoting. As a result, this athlete is susceptible to an anterior cruciate ligament (ACL) injury.

The second key equation to understand is:

Injury risk equation 2

Susceptible athlete + injury mechanism
= injury

Continuing the above example, should the netballer be exposed to an incident where she has to rapidly decelerate and change direction when catching and passing, all the ingredients are present for an ACL injury.

Each of these three variables (predisposition, extrinsic risk factors and injury mechanism) needs to be identified and either corrected or compensated for if we are to have an impact on reducing overall injury rates. The aim of screening, therefore, is to identify the predisposed individual so that the sum of these injury equations can be minimised.

Developing a screening tool

Step 1 – creation of the generic warning index

The first step in developing a pragmatic screening tool, therefore, is to make it as sport-/activity-specific as possible. In order to do this, it is vital that the injury profile of the sport be examined in detail. The sport's injury profile helps to form the basis of what can be termed a *warning index*. This warning index serves to highlight areas that need special attention paid to them. For example, as a rule, we would look closer at knee injury risk factors in skiers than we would in swimmers. This approach is based on epidemiological findings that tell us that knee injuries result in more competition and training time lost in skiing than in swimming.[2]

Epidemiological studies of injury rates in sport provide the basis for a pragmatic approach to injury risk identification. The governing bodies of many sports commission epidemiological studies into injury incidence and prevalence.

Most sports with a large participation base will have had some epidemiological study published in the scientific literature regarding injury rates over a specified period.

The two key factors to glean from these studies are:

(1) injury incidence, which is usually expressed as number of injuries reported per 1000 hours; and
(2) injury severity, which is usually expressed as games or training days missed.

> A good example is the Australian Football League's injury survey, which is the world's longest running publically released injury survey in sport. Other good examples are the annual injury reports produced by the Rugby Football Union (UK).

It is important to consider both factors because a relatively minor injury may occur frequently, whereas a less common injury may have effects that are more devastating and therefore warrant more investment in preventing it. An example of this is the ACL rupture that, whilst relatively infrequent, accounts for the most number of training days missed in a number of sports.[3] By combining both incidence and severity, a clearer insight into the top priorities for intervention can be gained.

It is vital that the comparison between these statistics and your athletes be valid, however. The sport itself may not be the only variable of interest. Injury patterns often differ depending on factors such as:

- Gender
- Competition level
- Age[4,5,6]

Accordingly, the more these statistics can be broken down the better. In doing this, a true profile of sport-specific injury incidence and severity can be determined, and benchmarks against the population average can be ascertained.

The downfall of using injury statistics is that often an injury is only recorded if an individual misses training or competing. In many cases, however, an injury may not stop an athlete from participation, although it may reduce their performance when doing so. Using the traditional injury surveillance format, these injuries do not count towards the total statistic and as a result may be under-represented. A more sophisticated method, therefore, is to report on injuries that affect performance, but at present this is not widely presented in the scientific literature.

If possible, it is even more powerful to be not just sport-specific, but also *position-specific*. This helps to account for the varying roles that different players within a team are expected to play, and therefore the varying loads to which they will be exposed. This often translates into a different injury profile. Examples of positional differences in injury profile can be seen in Table 6.1, below.

Once the injuries of greatest significance have been identified, the sport-specific generic warning index has been completed. This forms the basis for screening of the entire group of athletes that is to be tested. The next step is to customise the screen according to the intrinsic risk factors of each individual.

It is important to consider not just the demands of competition, but also the demands of training. In the ice hockey example, whilst ankle sprains do not feature as an injury of high significance, training often involves track running and plyometrics, both of which do have high incidences of ankle injuries.

By the end of this first step, we aim to identify the 4–5 major injuries that will form the heart of the assessment. The closer your group of athletes is to the groups found in the literature (based on factors such as age, gender, playing position, etc.), the more valid the comparison, which in turn increases the appropriateness of the list of injuries in the *generic warning index*.

Table 6.1 Examples of the impact of playing position on injury profile

Sport	*Positional differences in injury profile*
Rugby	Higher risk of shoulder instability in five-eight position compared to wing in elite rugby union.[7] A full exploration of injury differences according to playing position in professional rugby union players can be found in Brooks & Kemp (2011).
American football	Higher rates of shoulder injury in quarterbacks compared to tight ends.[8]
Rowing	Scullers tend to sustain more cervical injuries than sweep rowers.[9]

Step 2 – individualising the warning index

Athletes within a group will vary in injury risk profile according to their own intrinsic risk factors. Intrinsic risk factors are those internal to the individual, such as age, gender, race, body mass, muscle strength, physical fitness, biomechanics and history of injury. They can be divided into *modifiable* and *non*-modifiable factors. As the name suggests, non-modifiable factors are those that no manner of interventions can alter, such as age, gender or past injury profile. The presence in an individual of any of these non-modifiable risk factors should filter them into a specific testing pool. Modifiable risk factors will be dealt with in step 3.

The factor that has been shown to be the most powerful predictor of an injury is a past injury to that site. For example it is known the highest risk of a groin injury is a previous groin injury[10] and that individuals with a history of an ankle sprain have a higher chance of re-injuring that ankle.[11] This, therefore, is an area worthy of specific attention, even in a sport such as ice hockey where ankle sprains do not figure as an injury of significance.[12]

Injuries to other areas of the body may also provide a predictive clue to the risk of injuries to other body parts. For example a history of low back pain has been shown to be a powerful predictor of ACL injury in collegiate athletes.[13]

Another example of a targeted approach to screening is in the screening of females for ACL injury risk. Females are between 2–8 times more likely to sustain an ACL injury than males,[14] and it is our view that all females should be assessed for ACL injury risk in any land-based team sport.

By examining the profiles of the athletes or by conducting targeted interviews with them, the presence of many of these risk factors within the group can be ascertained prior to performing any tests. This helps direct the professional to a more targeted and individualised approach to injury screening.

Accordingly, the *interview questionnaire* is the next step in the process. Again, a myriad of factors could be screened for, but it is best to narrow the investigation to the areas of most interest according to the profile of the sport. An example can be seen in appendix 1. Again, it is not designed to be the definitive assessment tool, merely one that acts as a quick method of seeing if a specific injury (past or present) makes its way into the individual-specific warning index. It is a method of filtering information so that the areas of relevance can be distilled and then targeted in a physical examination.

The areas of interest as ascertained by the specific warning index can then be added to those in the generic warning index to give us intelligence-led targets to test during our physical examination. This demonstrates a real advantage of the warning index. Instead of having to examine the literature for risk factors for every single possible injury, our research can now be directed to the examination of those modifiable risk factors that relate specifically to the injuries of greatest significance to the individual.

> By the end of this second step, we aim to have identified both a generic and an individual-specific warning index. Such an intelligence-led approach enables us to develop a targeted and specific approach to risk factor identification, which increases its diagnostic power.

Step 3 – determining the risk factors

A myriad of factors have been proposed as being injury risk indicators, but to test for each of them is both unrealistic and unnecessary. The aim should be to make the screening tool as streamlined as possible, which is why we prioritise our modifiable risk factor research to those injuries that we have identified in steps 1 and 2.

In order to conduct the research, scientific databases (such as *Medline, sportdiscuss, Cinahl, Pubmed* and *Web of Knowledge*) need to be accessed with a variety of appropriate keywords entered. These keywords should be used as filters to ensure that the information most relevant to your group of athletes is attained. It is important not to filter too stringently; otherwise, important information may be missed. For example should the area of interest be concussion in lacrosse, it would be prudent to examine the literature regarding concussion in other contact sports as well.

Examples of key search words:

(name of body part), (name of injury), (sport name), child, youth, senior, amateur, professional, elite, sub-elite, male, female, risk factor, athletic, sport, injury, injuries, modifiable, intrinsic . . .

Table 6.2 below shows some factors that have been identified in previous research as being injury risk factors. These factors were identified using a string of words from the list of examples provided above.

It is vital to know these risk factors, as they direct the next stage of the process of screening tool development, which is to uncover the most valid and reliable assessments of athlete susceptibility to injury.

Once this third step has been completed, we should have generated a list of all the risk factors for the injuries that have been identified in the warning index.

Table 6.2 Examples of modifiable risk factors for sporting injuries

Injury	*Modifiable risk factor*
Groin injury	Hip internal rotation[15]
Hamstring injury	Hamstring muscle strength[16]
Lumbar spine stress fractures in cricket fast bowlers	Asymmetry of quadratus lumborum muscle bulk[17]
Concussion in sub-elite rugby league	Aerobic power[18]

Step 4 – selection of appropriate assessments of injury susceptibility

It is our belief that the reasons why pre-participation examinations have not been shown to have strong predictive validity[19] is that the screening tests selected are often (a) not specifically targeted at the athlete (therefore failing one of the first two stages) and/or (b) not valid or reliable.

The selection of appropriate tests is critical in the formulation of a screening tool. A number of things must be considered:

- Is the test itself a valid measure of the variable of interest?
- Does the test have sufficient inter- and intra-tester reliability?
- Is the *application* of the test valid?

Table 6.3 below describes these three issues. Ideally, all selected tests would have demonstrated reliability and validity, but it needs to be acknowledged that research often lags behind professional practice, and that clinical experience should not be discarded simply because it is yet to be scientifically proven. Nonetheless, it is reasonable to strive to include only the most appropriate tests. Accordingly, we propose a 3-tiered approach to test selection:

> *Tier 1*
> *Tests that achieve this status have been shown to be both valid and reliable, may have predictive value as injury predictors and are applicable in a clinical setting.*
>
> *Tier 2*
> *These tests have been found to be clinically useful and reliable but may be lacking in scientific proof of their validity or in their ability to be predictive of injury.*
>
> *Tier 3*
> *These tests may be clinically useful but may lack convincing scientific validity or reliability.*

Table 6.3 Essential criteria for test selection

Issue	*When present*	*When absent*
Test validity	The results of the test are an accurate indication of the variable being investigated.	The results of the test do not provide any true insight into the variable of interest.
Test reliability	The results are robust against random error and will be the same when repeated by the same or other people.	Different results may be obtained upon repeated testing by the same person. Equally, the results that one tester achieves may be different from that achieved by another tester. This means that results cannot be compared with confidence and so changes in the variable may not be 'true'.
Appropriate test application	This is a form of validity. It means that not only is the type of test selected appropriate, but the performance of the test gives a true indication of the variable of interest.	The test may be appropriate but the way in which it is performed may not give an insight into the variable of interest. For example isokinetic hamstring testing that does not examine its profile at high speeds (300°/sec).

The core of the screening assessment battery should be comprised of tests in tier 1 and tier 2. Tier 3 tests should be considered only if the high performance professional finds them particularly useful.

Other things that should be considered are the ease and expense of test application. This is particularly the case when there is a group of athletes to be tested on repeated occasions. The aim is to make the entire battery as specific, targeted and user-friendly as possible where the maximum amount of information can be gleaned with a minimum of time or money expenditure.

It is also important to point out that it is our view that all athletes that participate in contact or collision sports should undergo neurocognitive screening at the start of the season. This will serve to act as a baseline measure of higher-order cognitive functioning, one that can be used as a comparison should a concussion be sustained during the course of the season. This view is in keeping with best practice principles as outlined in the Zurich Consensus Statement on Concussion in Sport.[20]

Step 5 – assessment of movement proficiency

The fifth step requires the selection of tests that examine an individual's ability to perform specific functional tasks, which, in turn, is a display of their proficiency in integrating motion, stability and motor control. This step is essentially an extension of step 3, but we have given it a section of its own to emphasise the functional nature of these tasks.

The analysis of functional movement requires us to examine the system as a whole, which avoids a reductionist approach to examining parts as opposed to patterns. The background and therefore viewpoint of an interdisciplinary performance team is vital here, as is the use of video.

Analysing tasks such as landing control and gait analysis is vital, particularly for dynamic weight-bearing sports. Throwing or bowling technique is also crucial to examine in those

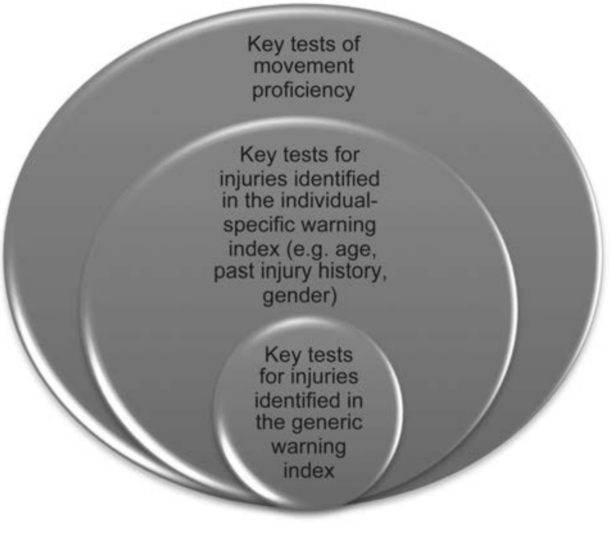


Figure 6.1 The 'onion skin' of risk factor identification

athletes that rely on such explosive upper limb tasks. Also included should be basic tasks, such as movement dissociation ability and squatting control, especially if loaded squatting will form part of a training regimen.

By this stage, the screening tool of musculoskeletal competencies should be complete. Depending on the level of detail required regarding an athlete, there may be some other pieces of information that can be integrated. Specifically, these may be a psychological profile as well as a medical screening, examining such things as cardiovascular and dental health.

The total screening tool, therefore, is an example of many members of the performance support team providing their expertise.

> Upon completion of the fifth step, we should have distilled the major intrinsic risk factors for an individual athlete based upon their risk index and movement profile. In essence, we will have developed and individualised an athletic profile and determined all the tests necessary to examine for the susceptible individual.

Step 6 – dealing with the results

The aim of the screening process is to assess risk and to highlight areas that require further investigation. For example a poor performance of an overhead squat, a reduction in shoulder ROM, or asymmetry when running or landing from a drop jump should alert members of the high-performance team to investigate further to ascertain the causes behind these poor results.

To illustrate this further, a poor squat may be due to a number of factors, including insufficient ankle or shoulder ROM, decreased lumbo-pelvic stability or poor technique/motor control. The sports medicine team can assess the relative contributions of ankle ROM to the performance of the squat, whereas the conditioning coach will be able to determine if their inability to squat is due to a technical deficiency. These issues can then be investigated further and managed appropriately.

This demonstrates the real value of the screen; it acts as a filter, whereby risk factors for injuries in the warning index deficits can be deconstructed into their individual parts. It also demonstrates how an interdisciplinary process will provide a thorough insight into the issues at hand. Frequently, a number of factors combine, and it is only through an integrated assessment that these factors can be evaluated and corrected. Again, the value of a multi-faceted screening process is evident.

The results of the screening should provide the high-performance staff with an accurate reflection of an individual athlete's deficiencies that need specific attention. These findings need to be discussed with the entire high-performance team. It is our opinion that this includes both the athlete and coaching staff. This is because all members of the team can provide some insight into the next part of the *injury risk equation,* which deals with the forces they are likely to be subjected to in the future. The marriage of these factors goes some way to providing the high-performance professional with a real insight into injury risk, which is, of course, the necessary first step in injury prevention. These modifiers provide the 'expected future forces' side of the injury risk equation.

Exposure to extrinsic factors known to increase the risk of a specific injury that the athlete is already predisposed to combine to turn the individual into a susceptible one. An example of such an interplay could be a female water polo player with a history of shoulder impingement pain and screening that has revealed anterior instability (predisposed athlete) about to enter a heavy preseason training period involving high-volume-resisted swimming and throwing (extrinsic risk factors). This athlete is now susceptible to further episodes of shoulder pain.

We would not know of her susceptibility to injury had we not understood the two parts of this injury equation. Clearly, all members of the high-performance team need to understand each part of the equation, as all have input into either increasing or decreasing her susceptibility to injury.

The bottom line is, however, that all findings of the screening process must be discussed with the performance and coaching teams and acted upon; otherwise, the entire process is futile. It must be viewed as the crucial first step in a strategic approach to injury risk reduction, but it is useless by itself.

> The completion of this step should provide all members of the interdisciplinary team with a plan of action for addressing any modifiable risk factors.

Step 7 – review

The screening tool is, by nature, ever-evolving because an individual's risk factors change with time, and the individual will enter or leave various risk groups according to these changes. This may be because during the previous season, they sustained a quadriceps tear, placing them in a risk group for sustaining another one, or they entered a 'high risk' age range (such as being over the age of 23 in the case of hamstring injuries).[21] Also, thanks to appropriate management, their hip ROM may have improved, reducing their risk of groin injuries.

As with any new process, there will be a 'bedding-in' period as the support staff gets used to performing the various assessment items. It is inevitable that some tests will be replaced by others over time due to professional preference or advancements in the scientific rationale for the risk factors of certain injuries. However, the philosophy of an intelligence-led, individually targeted approach to screening, using valid and reliable tests should remain constant.

How often should screening be performed?

The entire screening process should be completed at the start of each season. This gives the performance professional an indication of the status of a player before training load is applied. This also satisfies one of the other main objectives of screening, and that is to establish baselines, upon which deviations during the season can be measured.

It is considered prudent to repeat the screen in its entirety at the end of the season as well (with any necessary adjustments given any injuries sustained during the season). This can then be used as a means of prescribing off-season programmes. Dependent on time and resources, it could be argued that the process could be performed mid-season as well to ascertain any deviations from the preseason baseline.

It is becoming increasingly common in elite sport for an abridged screening process to be completed on a daily or weekly basis, particularly for those injuries of greatest concern, and using those tests that are most valid and sensitive to change. For example, in sports where groin injuries are of concern, or in an athlete with a recent history of groin injuries, a simple adductor squeeze test can be performed prior to training. It is known that a fall-off in adductor muscle strength precedes a groin injury, and so this test can form part of a quick daily or weekly screen.[22] Other examples may be an assessment of shoulder rotation strength using a handheld dynamometer as an indicator of shoulder function,[23] or a slump test to assess changes in neurodynamics. Fluctuations in these parameters can provide the performance professional with a clear indication of how the athlete is coping with training and competition load. Should there be a dramatic decline in these indices it can provide them with an evidence-based incentive to change programmes prior to an overuse injury occurring.

Summary

Musculoskeletal injury screening can be a very useful tool in the worthwhile endeavour to reduce injury burden in sport. We have presented a step-by-step framework for the development of an intelligence-led risk assessment tool. It is aimed at examining for the presence of the factors that should be of most concern to the performance professional. It does not, however, seek to be the definitive assessment tool for every injury. It is primarily a method of highlighting the areas of risk that require further attention paid to them. This method of filtering through the vast array of injuries that *could* occur to focus on the injuries that the athlete is most susceptible to will help inform effective management with the ultimate aim of injury risk reduction.

Case study

Eighteen-year-old female football (soccer) player in an elite development programme. Nil relevant medical history but a past injury history of a left anterior shoulder dislocation four months ago that resulted in three-month absence from competition following a period of conservative management.

Step 1: Generic warning index (GWI)

We have appraised the literature surrounding women's football. It has revealed the four injuries responsible for most time lost in the sport, which, along with concussion, forms the basis of our generic warning index.

Area to target	*Based upon*
Knee injury	Knee ligament ruptures are responsible for most training time lost of any injury in female football[24]
Lateral ankle sprain	Ankle sprains are the injury of highest incidence in female football[24]
Hamstring injury	Hamstring strains being the second-most prevalent injury in elite female football and the most common muscle strain[25]
Adductor-related groin pain	Groin muscle strains are the second most common muscular injuries seen in female football[26]
Neurocognition deficits	Aiming to set a baseline for neurocognitive competency in the event of a concussive event throughout the season

Step 2: Specific warning index (SWI)

Based upon our knowledge of the individual athlete's non-modifiable risk factors, we have been able to generate our specific warning index. For completeness, we have included ACL injury risk, despite the fact that it is found in the GWI.

Area to target	*Based upon*
Shoulder stability	Previous history of a shoulder dislocation. Very high recurrence rates (89%) have been demonstrated in individuals who sustained a traumatic anterior shoulder dislocation aged 30 or less[27]
Anterior cruciate ligament (ACL)	Increased risk of ACLI in female football players compared to their male counterparts[1]

Step 3: Selection of valid and reliable tests

These tests have been chosen as the most valid and reliable indicators of the risk factors identified for the injuries in the GWI and SWI.

Area of interest	*Test selected*	*Based upon*
Ankle	Knee-to-wall measurement of ankle dorsiflexion (DF) range of motion	Ankle DF ROM is a strong predictor of ankle sprain[28]
	Star excursion balance test	This test has been shown to be sensitive in detecting those with chronic ankle instability[29]
Knee	Isokinetic profile of hamstring strength at 300°/sec	Female athletes who have suffered an ACLI demonstrate hamstring weakness on isokinetic testing[30]
	Drop vertical jump	A valid and reliable method of assessing lower-limb biomechanics in female football (soccer) players. It has also been shown to be a predictor of ACL injury risk in female athletes[31]
Hamstrings	Isokinetic profile of eccentric hamstring strength	Shown to be positively correlated with hamstring muscle strains[32]
Adductors	Supine goniometric assessment of hip internal rotation (hip and knee positioned at 90°)	Shown to be a valid indicator of hip ROM, and reductions in internal rotation ROM has been demonstrated to be a powerful predictor of adductor injury risk[22]
	Adductor squeeze test	Found to be able to discriminate between athletes with and without groin pain[33]
	Active straight leg raise test	Shown to have diagnostic value when assessing adductor-related groin pain[34]

(*Continued*)

Area of interest	*Test selected*	*Based upon*
Anterior shoulder stability	Shoulder apprehension and relocation test	Highly specific and sensitive test of anterior shoulder instability[35]
Neurocognitive testing	Computerised neurocognitive or digit-symbol substitution test	Shown to be a valid and reliable measure of higher-order cognitive functioning following a concussive event[20]

Step 4: Assessment of movement proficiency

These tests have been selected to give an insight into motor control and kinematic competency in football-relevant tasks.

Test selected	*Rationale*
Landing technique in vertical drop jump	To quantify amount of hip and knee extension and knee valgus, all of which have been shown to be key factors in ACL injuries[31]
Analysis of 45° and 90° cutting technique	To quantify the amount of femoral internal rotation and adduction when cutting, factors implicated in ACL injury risk in female soccer players[1]
Overhead squat technique	To examine dynamic lumbar spine control and technique in a functional training task

Once this point is arrived at, we have developed a screening tool that is both sport-relevant and individualised to the specific athlete in question. Depending on the results of these tests, we then have an intelligence-led list of areas requiring in-depth assessment and management.

Notes

1 Alentorn-Geli, E., Myer, G. D., Silvers, H. J., Samitier, G., Romero, D., Lázaro-Haro, C., & Cugat, R. (2009). Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 1: Mechanisms of injury and underlying risk factors. *Knee Surgery Sports and Traumatology Arthroscopy Journal,* 17(7), 705–29.

2 Majewski, M., Susanne, H., & Klaus, S. (2009). Epidemiology of athletic knee injuries: A 10-year study. *Knee, 13*(3), 184–88.

3 Hägglund, M., Waldén, M., & Ekstrand, J. (2008). Injuries among male and female elite football players. *Scandinavian Journal of Medicine & Science in Sports, 19*(6), 819–27.

4 Frisch, A., Seil, R., Urhausen, A., Croisier, J. L., Lair, M. L., & Theisen, D. (2009). Analysis of sex-specific injury patterns and risk factors in young high-level athletes. *Scandinavian Journal of Medicine & Science in Sports, 19*(6), 834–41.

5 Gabbett, T. (2003). Incidence of injury in semi-professional rugby league players. *British Journal of Sports Medicine, 37*, 36–44.

6 Foreman, T., Addy, T., Baker, S., Burns, J., Hill, N., & Madden, T. (2006). Prospective studies into the causation of hamstring injuries in sport: A systematic review. *Physical Therapy in Sport*, *7*(2), 101–09.

7 Sundaram, A., Bokor, D. J., & Davidson, A. S. (2011). Rugby Union on-field position and its relationship to shoulder injury leading to anterior reconstruction for instability. *Journal of Science and Medicine in Sport, 14*, 111–14.

8 Kaplan, L. D., Flanigan, D. C., Norwig, J., Jost, P. & Bradley, J. (2005). Prevalence and variance of shoulder injuries in elite collegiate football players. *British Journal of Sports Medicine, 33*, 1142–46.

9 Wilson, F., Gissane, C., Gormley, J. & Simms, C. (2010). A 12-month prospective cohort study of injury in international rowers. *British Journal of Sports Medicine, 44*, 207–14.

10 Ryan, J., DeBurca, N., & McCreesh, K. (2014). Risk factors for groin/hip injuries in field-based sports: a systematic review. *British Journal of Sports Medicine, 48*, 1089–96.

11 Beynnon, B. D., Murphy, D. F., & Alosa, D. M. (2002). Predictive factors for lateral ankle sprains: A literature review. *Journal of Athletic Training*, *37*, 376–80.

12 Agel, J., Dompier, T., Dick, R., & Marshall, S. (2007). Descriptive epidemiology of collegiate men's ice hockey injuries: National Collegiate Athletic Association injury surveillance system, 1988–1989 through 2003–2004. *Journal of Athletic Training, 42*(2), 241–48.

13 Zazulak, B. T., Hewett, T. E., Reeves, N. P., Goldberg, B., & Cholewicki, J. (2007). Deficits in neuromuscular control of the trunk predict knee injury risk: A prospective biomechanical-epidemiologic study. *American Journal of Sports Medicine*, *35*(7), 1123–30.

14 Yu, B., & Garrett, W. E. (2007). Mechanisms of non-contact ACL injuries. *British Journal of Sports Medicine, 41*, 47–51.

15 Verrall, G. M., Slavotinek, J. P., Barnes, P. G., Esterman, A., Oakeshott, R. D. & Spriggins, A. J. (2007). Hip joint range of motion restriction precedes athletic chronic groin injury. *Journal of Science and Medicine in Sport, 10*, 463–66.

16 Opar, D. A., Williams, M. D., Timmins, R. G., Hickey, J., Duhig, S. J. & Shield, A. J. (2015). Eccentric hamstring strength and hamstring injury risk in Australian footballers. *Medicine and Science in Sport and Exercise, 67*, 857–65.

17 Ranson, C., Burnett, A., O'Sullivan, P., Batt, M. & Kerslake, R. (2008). The lumbar paraspinal muscle morphometry of fast bowlers in cricket. *Clinical Journal of Sports Medicine, 18*, 31–7.

18 Gabbett, T. J. & Domrow, N. (2005). Risk factors in subelite rugby league players. *American Journal of Sports Medicine, 33*, 428–34.

19 Peterson, A., & Bernhardt, D. (2011). The preparticipation sports evaluation. *Paediatric Respiratory Reviews, 32*(5), 53–7.

20 McCrory, P., Meeuwisse, W. H., Aubrey, M., Cantu, B., Dvo-ák, J., Echemendia, R. J., . . . Turner, M. (2014). Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine, 47*, 250–58.

21 Verrall, G., Slavotinek, G., Barnes, P., Fon, G., & Spriggins, A. (2001). Clinical risk factors for hamstring muscle strain injury: a prospective study with correlation of injury by magnetic resonance imaging. *British Journal of Sports Medicine, 35*(6), 435–39.

22 Verrall, G. M., Slavotinek, J. P., Barnes, P. G., Esterman, A., Oakeshott, R. D., Spriggins, A. J. (2007). Hip joint range of motion restriction precedes athletic chronic groin injury. *Journal of Science and Medicine in Sport*, *10*, 463–66.

23 Stickley, C. D., Hetzler, R. K., Freenyer, B. G., & Kimura, I. F. (2008). Isokinetic peak torque ratios and shoulder injury history in adolescent female volleyball athletes. *Journal of Athletic Training, 43*(6), 571–77.

24 Faude, O., Junge, A., Kindermann, W., & Dvorak, J. (2006). Risk factors for injuries in elite female soccer players. *British Journal of Sports Medicine, 40*, 785–90.

25 Jacobson, I., & Tegner, Y. (2007). Injuries among Swedish female elite football players: a prospective population study. *Scandinavian Journal of Medicine & Science in Sports, 17*(1), 84–91.

26 Tegnander, A., Olsen, O. E., Moholdt, T., Engebretsen, L., & Bahr, R. (2008). Injuries in Norwegian female elite soccer: a prospective one-season cohort study. *Knee Surgery Sports andTraumatology Arthroscopy Journal, 16*(2), 194–98.

27 Lill, H., Korner, J., Hepp, P., Verheyden, P., & Josten, C. (2001). Age-dependent prognosis following conservative treatment of traumatic shoulder dislocation. *European Journal of Trauma*, *27*(1), 29–33.

28 Wright, C. J., Arnold, B. L., Ross, S. E., Ketchum, J., Ericksen, J., & Pidcoe, P. (2013). Clinical examination results in individuals with functional ankle instability and ankle-sprain copers. *Journal of Athletic Training, 48*(5), 581–9.

29 Gribble, P. A., Hertel, J., & Plisky, P. (2012). Using the star excursion balance test to assess dynamic postural-control deficits and outcomes in lower extremity injury: A literature and systematic review. *Journal of Athletic Training, 47*(3), 339–57.

30 Myer, G. D., Ford, K. R., Barber Foss, K. D., Liu, C., Nick, T. G., & Hewett, T. E. (2009). The relationship of hamstrings and quadriceps strength to anterior cruciate ligament injury in female athletes. *Clinical Journal of Sports Medicine, 19*(1), 3–8.

31 Hewett, T., Myer, G. D., Ford, K. R., Heidt, R. S., Jr., Colosimo, A. J., McLean, S. G., . . . Paterno M. V. (2005). Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: A prospective study. *American Journal of Sports Medicine, 33*, 492–501.

32 Fousekis, K., Tsepis, E., Poulmedis, P., Athanasopoulos, S., & Vagenas, G. (2011). Intrinsic risk factors of non-contact quadriceps and hamstring strains in soccer: a prospective study of 100 professional players. *British Journal of Sports Medicine*, *45*, 709–14.

33 Malliaras, P., Hogan, A., Nawrocki, A., Crossley, K., & Schache, A. (2009). Hip flexibility and strength measures: reliability and association with athletic groin pain. *British Journal of Sports Medicine, 43,* 739–44.

34 Mens, J. M., Vleeming, A., Snijders, C. J., Koes, B. W., & Stam, J. (2002). Validity of the active straight leg raise test for measuring disease severity in patients with posterior pelvic pain after pregnancy. *Spine 27*(2), 196.

35 Lo, I. K., Nonweiler, B., Woolfrey, M., Litchfield, R., & Kirkley, A. (2004). An evaluation of the apprehension, relocation, and surprise tests for anterior shoulder instability. *American Journal of Sports Medicine, 32*(2), 301–307.

2

ANALYSIS OF DATA

Data is numerical information obtained by measuring or counting something. Statistics is a branch of applied mathematics that is used to analyze data so that the meaning of the data can be interpreted and evaluated. A statistical analysis involves the application of one or more statistical procedures (specific arithmetic calculations) to produce specific numbers that describe different characteristics of the data; each of these specific numbers is called a statistic. A statistical procedure may be lengthy, especially when there is a large amount of data. However, as most statistical procedures involve arithmetic operations that will be very familiar to all secondary school students (addition, subtraction, multiplication, division, square roots), the procedures will not be mathematically difficult. The purpose of this chapter is to introduce the fundamental concepts that underlie statistical analysis and to describe the statistics required to complete the worksheets that accompany Chapters 3 to 7.

Objectives

After reading this chapter, you should be able to do the following:

1 Distinguish continuous and discontinuous data.
2 Describe the four measurement scales.
3 Describe the normal distribution curve.
4 Describe the measures of central tendency.
5 Describe the measures of variability.
6 Calculate Pearson's product moment correlation coefficient.
7 Calculate a simple regression equation.

Continuous and discontinuous data

Most of the things that we measure (e.g. joint flexibility, muscle strength, leg power) or count (e.g. number of people who attend a sports centre each week, number of people who cycle to work each day) vary over time. Consequently, these phenomena and many other things that we measure or count are called variables. In general, there are two types of variables: continuous and discontinuous. Discontinuous variables are also referred to as discrete variables. The units of a continuous scale can be subdivided (to an extent that is dependent upon the precision of the measurement instrument), whereas the units of a discontinuous scale only occur as whole units, such as the number of spectators at a football game. Most things that can be measured yield continuous data. Most things that can be counted yield discontinuous data. Measurement is the process of comparing an unknown quantity, such as someone's weight, with some standard of the same variable, such as calibrated weighing scales, and then describing the previously unknown quantity in terms of the standard. The precision of a measurement is limited by the precision of the measurement instrument. For example, most bathroom scales measure weight to 0.1 kgf, but a more sophisticated force measurement system, such as a force platform in a biomechanics laboratory, may be able to measure weight to 0.01 kgf.

BOX 2.1

In general, there are two types of variables: continuous and discontinuous. The units of a continuous scale can be subdivided, whereas the units of a discontinuous scale only occur as whole units.

Measurement scales

Different types of data contain different amounts of information. For example, the numbers of students who passed and failed a particular exam is one type of data (number of passes, number of fails). Analysis of pass/fail data is restricted to calculating percentages, in this case, the percentage of students in the group that passed the exam and the percentage of students that failed. In comparison, the actual marks obtained by the students in the exam is a different type of data which not only allows the calculation of percentage pass and percentage fail, but also allows the calculation of a number of other statistics to provide a clearer interpretation and more in-depth evaluation of the performance of the students in the exam. Consequently, the amount of information contained in a set of data depends upon the measurement scale used in the collection of the data. There are four generally accepted measurement scales (also referred to as measurement levels): nominal, ordinal, interval and ratio (Stevens 1946).

A nominal scale consists of a number of mutually exclusive categories that are qualitatively different with respect to a particular variable. For example, a nominal scale could be developed for hair colour with categories such as black, brown, blonde, red and grey. Similarly, a nominal scale could be developed for customer satisfaction with several categories ranging from very satisfied to very dissatisfied. The order of the categories in a nominal scale is not relevant as there is no quantitative difference between the categories. For example, it makes no sense to suggest that black hair is quantitatively different to brown hair. Category membership is the only information in a nominal scale of measurement

An ordinal scale, also referred to as a rank-order scale, is one in which the data can be ordered (or ranked) from low to high (or high to low) with respect to a particular variable without any knowledge of the actual differences between the ranks. For example, a group of soldiers could be arranged in order of increasing height with the number 1 assigned to the shortest soldier, the number 2 assigned to the next shortest soldier, and so on, with the highest number (rank) being assigned to the tallest soldier. When all of the soldiers had been given a rank-order number, the assigned numbers would represent an ordinal scale. By knowing the rank assigned to a particular soldier, we would not know the actual height of the soldier nor by how much he was taller or shorter than the soldiers adjacent to him. However, we would know whether he was shorter or taller than any other soldier in the group. Similarly, each runner that completes a marathon race is normally given a number that indicates the order in which she finished the race. The list of numbers assigned to the runners, from first to last, represents an ordinal scale. Each number indicates the position in which a particular runner finished the race, but not the amount of time between the runner and the runner who finished just in front or the runner who finished just behind.

In an interval scale, the data can be ranked and there is uniformity of difference between the units in the scale such that the actual difference between ranks can be accurately quantified. However, in an interval scale there is no zero. For example, temperature is measured on an interval scale where equal distances on the temperature scale indicate equal changes in heat. However, it is incorrect to say that the temperature at any particular point A on the scale is so many times hotter or colder than that at some other point B because the points A and B are not measured from true zero. The Celsius scale of temperature defines the freezing point of water as zero degrees (0 ºC) and the boiling point of water as 100 degrees (100 ºC), but the Celsius scale, satisfactory for measuring temperature in daily living, represents only part of the global temperature scale which extends below 0 ºC and above 100 ºC.

A ratio scale is the most useful measurement scale as it includes all of the characteristics of the other three scales and, in addition, has a true zero. Time is a ratio variable: 4 minutes is twice as long as 2 minutes. Age, height, weight, distance, speed, acceleration, force and power are also ratio variables.

BOX 2.2

There are four generally accepted measurement scales: nominal, ordinal, interval, ratio.

Normal distribution curve

The word score is used in different ways which include, for example, a test score, a golf score and a credit score. In statistics, a score is that which results from any act of measuring or counting. Any score taken directly, without alteration, such as measurement of body weight on a set of weighing scales, is called a raw score. In statistics, a raw score is usually denoted by the capital letter *X* and a collection of raw scores is often referred to as a distribution of scores. The total number of scores in a distribution of scores is usually denoted by the capital letter *N*. When *N* is large and the variable is discontinuous, it is helpful to start the process of data analysis by organizing the scores into a manageable form by (i) identifying the lowest score and the highest score in the distribution of scores, (ii) listing, in ascending order, each possible score between and including the lowest and highest scores, (iii) counting the frequency of each actual score in the distribution. The resulting distribution of frequencies is referred to as a frequency distribution. Table 2.1 shows a frequency distribution based on the heights (to the nearest whole centimetre) of 247 adult male applicants for entry to a physical fitness instructor course. The sum of the frequencies should be the same as the total number of scores, in this case, $N = 247$. In this frequency distribution the lowest score is 164 cm (5′ 4 1/2″) and the highest score is 194 cm (6′ 4 1/2″). The difference between the highest score and the lowest score in a distribution is called the range, usually denoted in statistics by the capital letter *R*. The range in the distribution in Table 2.1 is 30 cm ($R = 194$ cm − 164 cm = 30 cm). Inspection of Table 2.1 indicates that the frequency of scores tends to increase from the lowest score toward the middle of the range and then decrease from the middle of the range toward the highest score. The trend is shown more clearly in Figure 2.1a which shows the frequencies plotted against the scores. In a two-dimensional graph such as Figure 2.1a, it is customary to refer to the vertical axis as the ordinate (or *y* axis) and the horizontal axis as the abscissa (or *x* axis). The roughly bell-shaped frequency distribution of scores for height in Figure 2.1a is typical of the frequency distribution of many other variables concerning body size and shape; these include body weight, leg length, waist girth, hat size, shirt size and shoe size. For these and other variables, such as IQ (intelligence quotient), the greater the number of scores in the distribution, the closer the frequency distribution tends to approximate a symmetric bell-shaped curve, referred to as the normal distribution curve, normal probability curve and Gaussian curve, as shown in Figure 2.1b. Any distribution of scores which closely approximates the normal probability curve is said to be normally distributed. Figure 2.1c shows the similarity between the frequency distribution in Figure 2.1a and the normal distribution curve.

BOX 2.3

When the total number of scores is large, many variables concerning body size and shape are normally distributed.

TABLE 2.1 Frequency distribution of the heights of 247 men. *X*, raw scores; *f*, frequency of each score; *Xf*, product of each score and the frequency of each score; *CF*, cumulative frequency of the scores from the lowest score to the highest score; $\sum Xf$, sum of all scores.

X (cm)	*f* (n)	*Xf* (cm)	*CF* (n)
194	1	194	247
193	0	0	246
192	1	192	246
191	0	0	245
190	1	190	245
189	3	567	244
188	2	376	241
187	4	748	239
186	6	1116	235
185	6	1110	229
184	10	1840	223
183	15	2745	213
182	18	3276	198
181	21	3801	180
180	22	3960	159
179	24	4296	137
178	22	3916	113
177	19	3363	91
176	20	3520	72
175	14	2450	52
174	9	1566	38
173	8	1384	29
172	6	1032	21
171	5	855	15
170	3	510	10
169	3	507	7
168	2	336	4
167	0	0	2
166	1	166	2
165	0	0	1
164	1	164	1

$N = 247$

$\sum Xf = 44180$

$\overline{X} = \text{Mean} = \text{sum } Xf/\text{N} = 44180/247 = 178.87$ cm

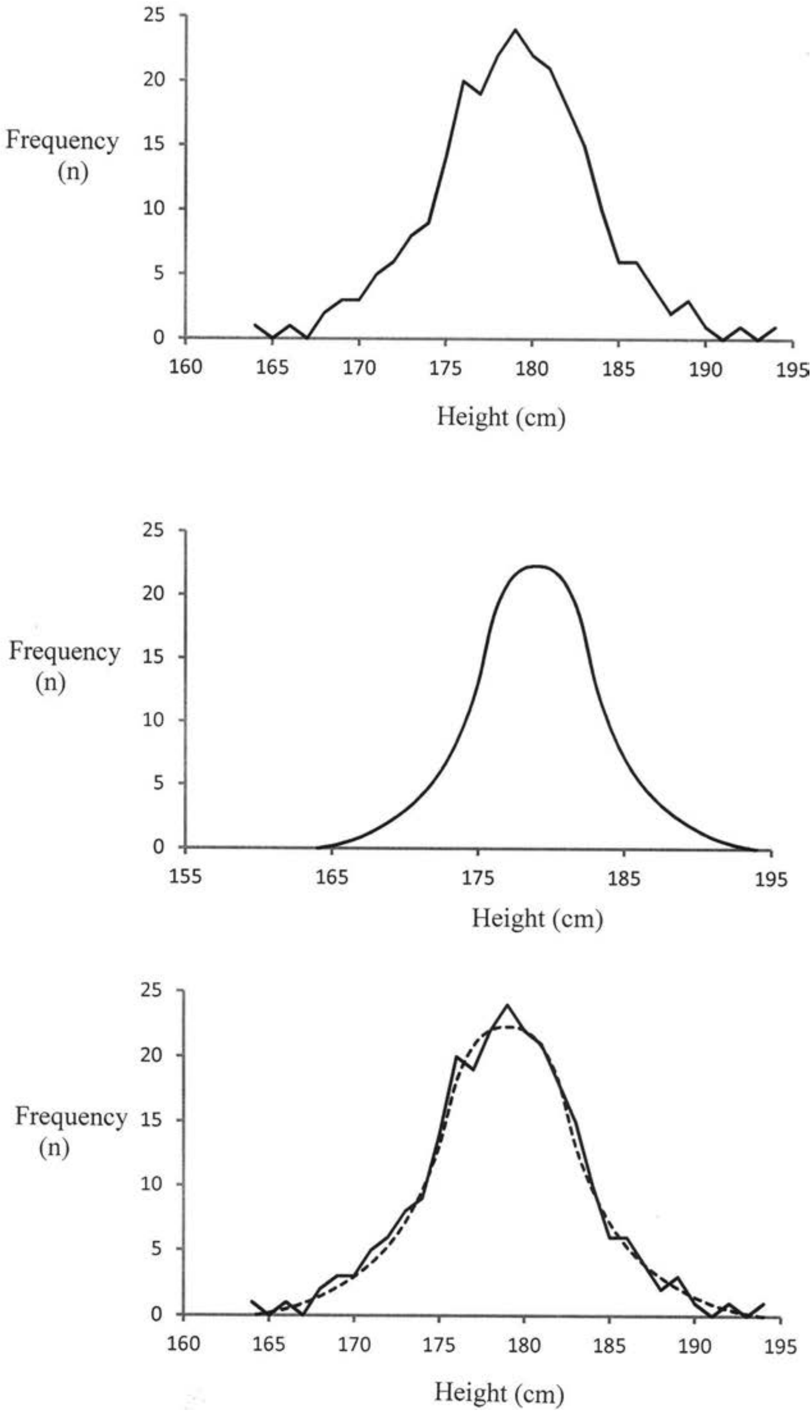


FIGURE 2.1 (a) Frequency distribution of the height of 247 men. (b) Normal distribution curve. (c) Normal distribution curve superimposed on the frequency distribution.

Measures of central tendency

As most of the scores in a distribution tend to be around the middle of the distribution, it is very common for a mid-range score to be used to reflect or represent the distribution as a whole. For example, a student may describe her whole-year academic performance in terms of an average grade or average percentage. Similarly, the socio-economy of a particular region might be described in terms of the average wage or average house price. In statistics, these measures are known as measures of central tendency. There are three measures of central tendency: mean, median and mode.

The mean, also referred to as the arithmetic mean and average, is the most useful and most widely used measure of central tendency. The mean of a distribution is calculated by dividing the sum of the scores by the number of scores, i.e.

$$\overline{X} = \frac{\sum X}{N}$$

where
$\overline{X}$ = the mean ($\overline{X}$ is read as X bar)
$\sum X$ = the sum of X (the sum of the scores)
N = the number of scores

The mean of the distribution 2, 5, 7, 5, 8, 5, 6 is 5.43, i.e.

$$\overline{X} = \frac{2+5+7+5+8+5+6}{7} = \frac{38}{7} = 5.43 \text{ (to two places of decimals)}$$

The mode of any distribution is the score that occurs most frequently. In the distribution of scores consisting of 2, 5, 7, 5, 8, 5, 6 the mode is 5 (frequency = 3). In the distribution shown in Table 2.1, the mode is 179 cm (5′ 10 1/2″) (frequency = 24). The mode tends to be the most useful measure of central tendency in the manufacture of certain goods such as clothing and footwear. For example, trouser manufacturers produce trousers to fit particular waist girth/leg length ranges and production will be geared to the number of sales of the different ranges. As most people will be of intermediate size, most of the trousers produced will be in the intermediate range. Similarly, shoe manufacturers produce shoes to fit particular length/width ranges. As most people will be of intermediate size, most of the shoes produced will be in the intermediate range. When a distribution has one mode, it is referred to as a unimodal distribution. Some distributions may have more than one mode, i.e. two or more scores may exhibit distinct peaks in the distribution. If there are two such scores, the distribution is referred to as bimodal, and if there are three such scores, the distribution is referred to as trimodal.

The median score in a distribution is the middle score in the distribution when the distribution is listed from lowest score to highest score. In the distribution of scores consisting of 2, 5, 7, 5, 8, 5, 6, the order becomes 2, 5, 5, 5, 6, 7, 8 when the scores are listed from low to high. As there are seven scores in the distribution, the median is the fourth score in the list, i.e. 5. When there is an even number of scores in a distribution, the median is the mean of the middle two scores. In the distribution 2, 3, 5, 6, 8, 9, the median is 5.5 ((5 + 6)/2 = 5.5). The median is useful when there are one or more outliers in the distribution, i.e. scores that are much higher or much lower than all of the other scores in the distribution and, if included in the calculation of the mean, would be likely to result in the data being misinterpreted. For example, in the distribution 2, 3, 4, 5, 6, 7, 23, the highest score, 23, is much higher than all of the other scores. The mean of the distribution is 7.14 (2 + 3 + 4 + 5 + 6 + 7 + 23 = 50; 50/7 = 7.14), but the median is 5. The median score of 5 is much more representative of the majority of the scores than the mean score of 7.14. Outliers may occur due to measurement error or to experimental error.

For ratio data, the mean is the best measure of central tendency because it is based on all of the scores in the distribution. In contrast, the mode is simply the single score that occurs most frequently and the median is a single score or the mean of only two scores in the distribution. When *N* is large, the distribution of scores for many human attributes including, for example, height and weight, is likely to be close to normal, i.e. the mean, median and mode will be close to each other. For example, in the distribution of the heights of 247 adult men shown in Table 2.1, the mean = 178.87 cm, the mode = 179 cm (frequency = 24) and the median = 179 cm (124th score). In a perfectly normal distribution, the mean, mode and median will be exactly the same. The most common type of deviation from a normal distribution is a skewed distribution, i.e. one in which there are a disproportionate number of scores close to one end of the distribution. Figure 2.2a shows a distribution in which most of the scores are in the lower part of the distribution. This type of distribution is described as positively skewed. A positively skewed distribution may reflect, for example, the results of an exam in which many of the students obtained a low score. In a positively skewed distribution, the mode will be lower than the median and the median will be lower than the mean. Figure 2.2b shows a distribution in which most of the scores are in the upper part of the distribution. This type of distribution is described as negatively skewed and may reflect, for example, the results of an exam in which many of the students obtained a high score. In a negatively skewed distribution, the mode will be higher than the median and the median will be higher than the mean.

BOX 2.4

There are three measures of central tendency: mean, median and mode.

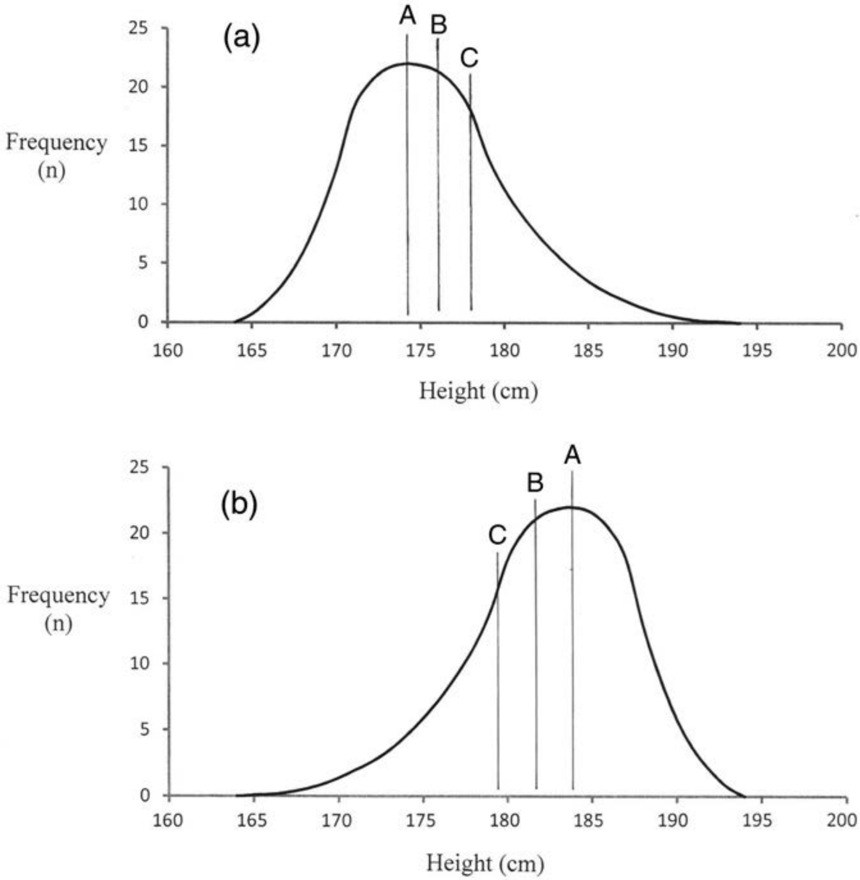


FIGURE 2.2 (a) A positively skewed distribution. (b) A negatively skewed distribution. A: mode; B: median; C: mean.

Measures of variability

A measure of central tendency provides useful but limited information about a particular distribution. For example, the average temperature in London, England, is 10.4 ºC (50.8 ºF), but the temperature ranges between a low of –3 ºC (26.6 ºF) and a high of 33 ºC (91.4 ºF) (Weather Online 2016). Clearly, in addition to a measure of central tendency, it would be helpful to have a measure that describes the spread or dispersion of the scores with respect to the central tendency. There are three statistics that are commonly used to indicate the spread or dispersion of scores in a distribution: range, average deviation and standard deviation. These statistics are known as measures of variability.

The range is the simplest measure of variability, but as it is based on only two scores, the lowest score and the highest score, it gives no indication of the spread of the other scores about the central tendency. The average deviation and standard deviation are both based on all of the scores in a distribution. The difference between any score in a distribution and the mean of the distribution is called the score's deviation. The deviation is symbolized by the lower case letter x and $x = X - \overline{X}$, i.e. the mean is subtracted from the score so that scores larger than the mean are associated with positive deviations and scores smaller than the mean are associated with negative deviations. Table 2.2 shows a distribution of scores consisting of the weights of 12 men. The mean of the distribution is 82.5 kgf as shown in the second column of Table 2.2. The deviations of the scores are shown in the third column of Table 2.2 together with the sum of the deviations, $\sum x$, which is zero. In any distribution of scores, irrespective of the size of N, the sum of the positive deviations (+44.5 in Table 2.2) must equal the sum of the negative deviations (–44.5 in Table 2.2) and, therefore, the sum of the deviations must be zero. Consequently, it is not possible to calculate the average deviation based on the positive and negative deviations because the numerator of the formula for average deviation, $\sum x/N$, will always be zero. However, it is possible to calculate the average deviation by treating all of the deviations as absolute deviations, i.e. all deviations are considered positive, irrespective of whether a score is above or below the mean. For example, a deviation of 1.5 kgf below the mean is the same distance from the mean as a deviation of 1.5 kgf above the mean. When the signs of the deviations are ignored, the deviations become absolute deviations. The mathematical notation for an absolute value is $|n|$, where n is the number whose absolute value is to be taken. Thus, $|-5| = |5| = 5$. The average deviation must be calculated from the absolute deviations, i.e.

$$\text{Average deviation} = \frac{\sum|x|}{N}$$

where

$\sum|x|$ = the sum of the absolute deviations

N = the number of scores

The fourth column of Table 2.2 shows the absolute deviations. The sum of the absolute deviations is 89, which results in an average deviation of 7.42 kgf, i.e.

$$\text{Average deviation} = \frac{\sum|x|}{N} = \frac{89}{12} = 7.42\,\text{kgf}$$

$$\overline{X} = 82.5 \quad \sum x = 0 \quad \sum|x| = 89 \quad \sum x^2 = 1155$$

TABLE 2.2 Average deviation, variance and standard deviation of the weights (X) of 12 men. AD = average deviation; VAR = variance; SD = standard deviation.

Score	X (kgf)	$x = X - \overline{X}$ (kgf)	$\lvert x \rvert$ (kgf)	x^2 (kgf^2)
1	72	–10.5	10.5	110.25
2	85	2.5	2.5	6.25
3	77	–5.5	5.5	30.25
4	67	–15.5	15.5	240.25
5	82	–0.5	0.5	0.25
6	82	–0.5	0.5	0.25
7	97	14.5	14.5	210.25
8	102	19.5	19.5	380.25
9	89	6.5	6.5	42.25
10	84	1.5	1.5	2.25
11	82	–0.5	0.5	0.25
12	71	–11.5	11.5	132.25

$$AD = \frac{\sum|x|}{N} = \frac{89}{12} = 7.42\,\text{kgf}$$

$$VAR = \frac{\sum x^2}{N} = \frac{1155}{12} = 96.25\,\text{kgf}^2$$

$$SD = \sqrt{\frac{\sum x^2}{N}} = \sqrt{96.25} = 9.81\,\text{kgf}$$

An increase or decrease in the spread of scores will be reflected in a corresponding increase or decrease in the average deviation. Like the average deviation, the standard deviation will also increase or decrease as a result of an increase or decrease in the spread of scores. However, unlike the average deviation, the standard deviation can be used to calculate other statistics for distributions of scores that approximate normal distributions; these statistics provide a more detailed analysis of the data. As many variables concerning human characteristics and behaviours are normally distributed (as in Figure 2.1b) when N is large, the standard deviation is the most commonly reported measure of variability.

The normal distribution curve is a regular geometric figure, i.e. the symmetric bell-shaped relationship between frequency and score is embodied in a mathematical equation (not required in this book). The slope of the normal distribution curve constantly changes (Figure 2.3a). The slope is close to zero (almost parallel with the abscissa) at the low end of the range and then increases gradually to a point of maximum steepness called the point of inflection. Between the point of inflection and the peak of the curve the slope gradually decreases to zero. Between the peak of the curve and the high end of the range, the curve is a mirror image of that between the peak and the low end of the range. In a normal distribution, the distance along the abscissa between the mean of the distribution, $\overline{X}$ (the peak of the curve where the slope is zero) and the point of inflection (where the slope is steepest) is called the standard deviation (Figure 2.3a). The standard deviation, denoted in statistics by σ, the lower case Greek letter sigma, is a key element in the mathematical equation of the normal distribution curve.

The area between the normal distribution curve and the abscissa represents all of the scores in a distribution. If the standard deviation of the distribution is known, it is possible to determine the percentage of scores contained in any section of the area with respect to the abscissa. If the curve is normal, the area between the mean $\overline{X}$ and $\overline{X} + 1\sigma$ is always 34.13% of the total area, i.e. for any normally distributed set of scores, 34.13% of the scores will lie between $\overline{X}$ and $\overline{X} + 1\sigma$ (Figure 2.3b). Similarly,

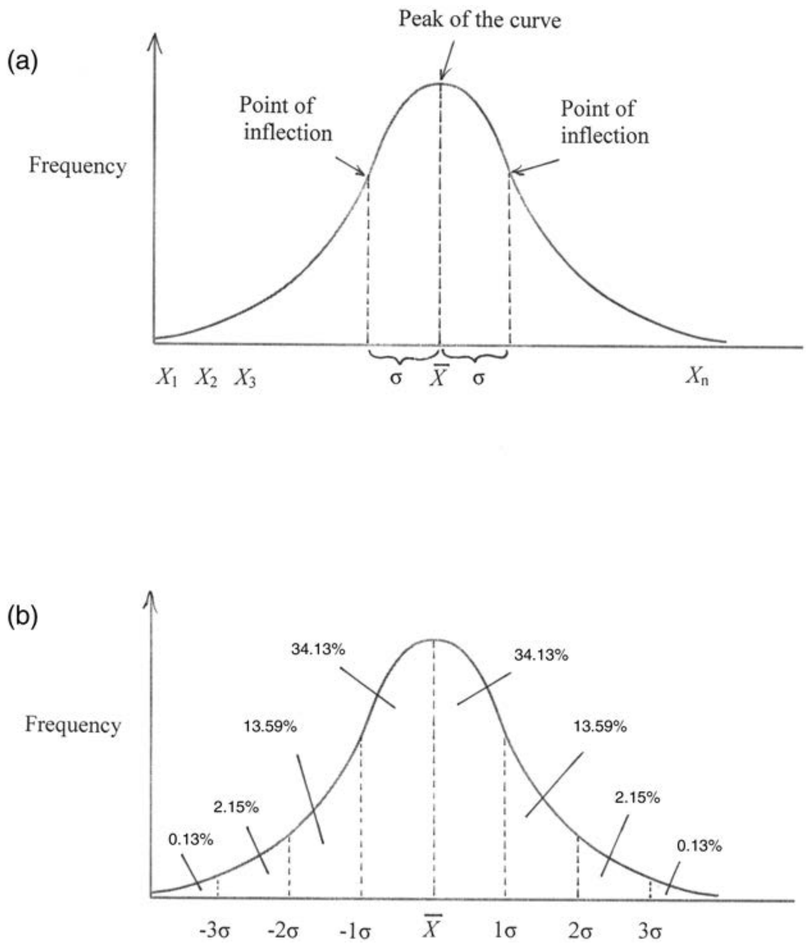


FIGURE 2.3 (a) The relationship between the mean and the standard deviation in the normal distribution curve. (b) The areas under the normal distribution curve defined by standard deviation units.

34.13% of the scores will lie between $\overline{X}$ and $\overline{X} - 1\sigma$. Standard deviations are always measured from the mean. The percentage of the total area between $+1\sigma$ and $+2\sigma$ is 13.59%, as is the percentage of the total area between -1σ and -2σ. The percentage of the total area between $+2\sigma$ and $+3\sigma$ is 2.15%, as is the percentage of the total area between -2σ and -3σ. For most practical purposes, the area contained within three standard deviation units in both directions (99.74% of the total area) can be considered to include 100% of the scores.

As described earlier, the average deviation is the mean of the sum of the absolute deviations ($\sum|x|/N$). The standard deviation is the square root of the mean of the sum of the squared deviations, i.e.

$$\text{Standard deviation} = \sigma = \sqrt{\frac{\sum x^2}{N}}$$

where
$\sum x^2$ = the sum of the squared deviations
N = the number of scores

The fifth column of Table 2.2 shows the squared deviations together with the sum of the squared deviations (1155 kgf^2). Table 2.2 also shows the mean of the sum of the squared deviations (96.26 kgf^2) and the standard deviation (9.81 kgf), i.e.

$$\sum x^2 = 1155\ \text{kgf}^2$$

$$\frac{\sum x^2}{N} = \sigma^2 = \frac{1155\,\text{kgf}^2}{12} = 96.26\,\text{kgf}^2$$

$$\sqrt{\frac{\sum x^2}{N}} = \sigma = \sqrt{(96.26\,\text{kgf}^2)} = 9.81\,\text{kgf}$$

The mean of the sum of the squared deviations ($\sum x^2/N = \sigma^2 = 96.26\ \text{kgf}^2$) is called the variance. The variance is an important statistic in detailed analyses of normally distributed data.

Some authorities maintain that the denominator in the equation for the calculation of the standard deviation should be $N-1$ rather than N. Unless N is very small, the use of $N-1$ makes very little difference to the value of the standard deviation or to interpretation of the standard deviation. Using $N-1$ is likely to give a better estimate of the standard deviation of the population from which the sample of N scores was taken (Spiegel *et al.* 2013). However, in most circumstances the user is only interested in describing the distribution of N scores, in which case, the use of N is preferred. In the distribution of scores in Table 2.2, $\sigma = 9.81$ kgf when N is used and $\sigma = 10.2$ kgf when $N-1$ is used.

The average deviation is always smaller than the standard deviation. In a perfectly normal distribution the average deviation is always 0.79788σ (Spiegel *et al.* 2013). For the distribution of scores in Table 2.2, the average deviation is 7.42 kgf and the standard deviation is 9.81 kgf, i.e. the average deviation = 0.756σ (7.42/9.81 = 0.756) which, not surprisingly, indicates that the distribution of 12 scores is not normal. For the distribution of scores in Table 2.1, the average deviation is 3.59 cm and the standard deviation is 4.63 cm, i.e. the average deviation = 0.775σ (3.59/4.63 = 0.775), which indicates that the distribution of 247 scores is not normal, but closer to normal than the much smaller distribution of 12 scores in Table 2.2.

In a normal distribution, 68.26% of scores would be expected to lie between −1σ and +1σ (Figure 2.3b). The mean and standard deviation of the distribution of 247 scores in Table 2.1 are 178.87 cm and 4.63 cm, respectively. Consequently, the −1σ to +1σ range corresponds to 178.87 ± 4.63 cm, i.e. 174.24 cm to 183.5 cm. Inspection of the first column of Table 2.1 shows that 74.5% of the scores (184 of the 247 scores: 184/247 × 100 = 74.5%) lie within the 174 cm to 183 cm range, which indicates that the distribution approximates a normal distribution (as shown in Figure 2.1c).

BOX 2.5

There are three measures of variability: range, average deviation and standard deviation.

Reliability

The usefulness of any test or measurement is determined by its validity and reliability. Validity refers to the characteristic or attribute that the measurement is supposed to assess. For example, a valid test of muscle strength will assess strength rather than, for example, muscular endurance. Similarly, a valid test of power will assess power rather than, for example, speed. Validity is covered in the next main section. Reliability refers to the consistency of a measurement, i.e. the stability of a score in repeated measurements. For example, if the height and weight of each member of a group of people are carefully measured by trained personnel using properly calibrated equipment and then measured again a few minutes later by the same personnel using the same equipment, it is likely that the second set of measures will be exactly the same as the first set of measures. Consequently, height and weight are very reliable measures when the amount of time between repeated measures is short. In adults, height is a very reliable measure over much longer periods of time as diurnal bodily changes are unlikely to affect height. However, diurnal bodily changes are likely to affect weight. Consequently, weight is a less reliable measure than height in adults. Even so, for most adults, height, weight and other anthropometric measures tend to be very reliable measures over long periods. In contrast, measures of attributes that involve considerable mental and physical effort as in, for example, tests of strength, endurance, speed and power, may vary considerably for individuals and between individuals over quite short periods due to differences in habituation, learning, motivation, fatigue and other diurnal variations. A test or measure cannot be considered valid if is it not reliable. A test or measure can be reliable but not valid.

Correlation

The reliability of a test or measure is determined by correlation, i.e. a statistical technique which expresses the strength of the relationship between two sets of scores. In this context, relationship refers to the level of agreement in the rankings of the scores and in the differences between the ranks in repeated measures; the greater the similarity in the rankings of the scores and in the differences between the ranks, the greater the reliability of the measurement. The strength of the relationship between two sets of scores is expressed numerically by a statistic called the coefficient of correlation. The coefficient of correlation ranges between +1 and –1, where +1 indicates a perfect positive correlation (an increase in score in one set of scores is associated with an equivalent increase in score in the second set of scores), zero indicates no relationship and –1 indicates a perfect negative relationship (an increase in score in one set of scores is associated with an equivalent decrease in score in the second set of scores). The most common method of calculating the coefficient of correlation is the Pearson Product Moment method (after the English statistician Karl Pearson, 1857–1936). The coefficient of correlation obtained by using the Pearson Production Moment method is the statistic referred to as Pearson's *r* (lower case letter *r*). Figure 2.4 shows

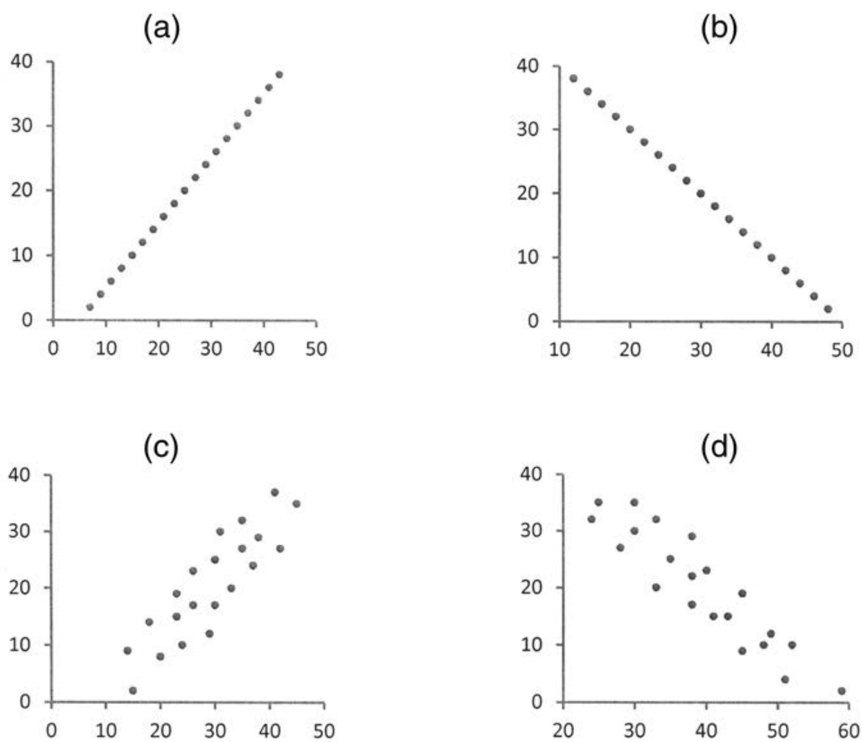


FIGURE 2.4 Scattergrams to illustrate different levels of correlation. (a) Perfect positive correlation, $r = +1$. (b) Perfect negative correlation, $r = -1$. (c) Positive correlation, $r = +0.88$. (d) Negative correlation, $r = -0.92$.

four scattergrams which illustrate different levels of correlation. In a scattergram, each pair of scores (first score on the x axis, second score on the y axis) is plotted as a single point. Figure 2.4a shows a perfect positive relationship ($r = +1$). Figure 2.4b shows a perfect negative relationship ($r = -1$). Figure 2.4c shows a positive relationship of $r = +0.88$ and Figure 2.4d shows a negative relationship of $r = -0.92$. The two sets of data used to plot Figure 2.4c are shown in Table 2.3 together with the calculation of r.

BOX 2.6

The strength of the relationship between two sets of scores is expressed numerically by a coefficient of correlation which ranges between –1 and +1. The most commonly reported coefficient of correlation is Pearson's r.

TABLE 2.3 Calculation of the Pearson Product Moment coefficient of correlation between two variables, X and Y. The number of scores in each set of scores is 21.

Score	X	x	x^2	Y	y	y^2	xy
1	15	–14.29	204.08	2	–18.57	344.90	265.30
2	14	–15.29	233.65	9	–11.57	133.90	176.88
3	20	–9.29	86.22	8	–12.57	158.04	116.73
4	18	–11.29	127.37	14	–6.57	43.18	74.16
5	24	–5.29	27.94	10	–10.57	111.75	55.88
6	23	–6.29	39.51	19	–1.57	2.47	9.88
7	23	–6.29	39.51	15	–5.57	31.04	35.02
8	29	–0.29	0.08	12	–8.57	73.47	2.45
9	26	–3.29	10.79	23	2.43	5.90	–7.98
10	26	–3.29	10.79	17	–3.57	12.75	11.73
11	30	0.71	0.51	17	–3.57	12.75	–2.55
12	30	0.71	0.51	25	4.43	19.61	3.16
13	33	3.71	13.80	20	–0.57	0.33	–2.12
14	37	7.71	59.51	24	3.43	11.75	26.45
15	31	1.71	2.94	30	9.43	88.90	16.16
16	35	5.71	32.65	27	6.43	41.33	36.73
17	38	8.71	75.94	29	8.43	71.04	73.45
18	32	2.71	7.37	27	6.43	41.32	17.45
19	41	11.71	137.22	37	16.43	269.90	192.45
20	45	15.71	246.94	32	11.43	130.61	179.58
21	45	15.71	246.94	35	14.43	208.18	226.74
			$\sum x^2 = 1604.29$			$\sum y^2 = 1813.14$	$\sum xy = 1507.57$

$$x = X - \overline{X}$$

$$y = Y - \overline{Y}$$

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} = \frac{1507.57}{\sqrt{(1604.29)(1813.14)}} = \frac{1507.57}{1705.52} = 0.88$$

The most common method of assessing the reliability of a test or measure is to calculate r from two sets of scores obtained by administering the test on two occasions. This type of reliability is referred to as 'test–retest reliability' and the resulting correlation coefficient is referred to as the test's 'reliability coefficient'. In tests and measures of physical fitness, the time between test and retest should allow adequate recovery, but be short enough to prevent the second set of scores being affected by growth and development. Two or three days of rest may be required to ensure adequate recovery from some physiological tests that involve prolonged intense physical effort. However, a few minutes of rest should ensure adequate recovery from tests that are completed fairly quickly, such as single jumps or short sprints.

Standard error of measurement

Table 2.4 shows the test and retest scores for a test of peak instantaneous mechanical power output in a countermovement vertical jump with hands on hips for a group of 26 male rugby players. After a few practice trials (practice jumps) to familiarize themselves with the movement, the players performed two test trials with 5 minutes' rest between the trials. The reliability coefficient for the test was found to be $r = 0.9255$ (the plus sign is usually omitted for positive correlations). Figure 2.5 shows a scattergram of the test (x axis) and retest (y axis) scores. A reliability coefficient of $r = 0.9255$ is considered very high in the classification of Guilford (1950) (see Table 2.5) and would suggest that the test is a reliable test of peak instantaneous power in male rugby players. However, the reliability coefficient of a test does not indicate whether the test is reliable enough for its intended purpose. For example, if the test of peak instantaneous power is to be used to monitor the effects of a power training programme with a group of male rugby players, some estimate of the precision of measurement afforded by the test is required so that changes in peak instantaneous power that occur following training can be properly interpreted as real changes or measurement error. A statistic called the standard error of measurement (*SEM*) of a test provides an estimate of the precision of measurement (Thomas *et al.* 2005). The $SEM = s\sqrt{(1 - r)}$, where s is the standard deviation of all of the test scores and retest scores and r is the reliability coefficient. For the

TABLE 2.4 Reliability coefficient (r) and standard error of measurement (SEM) for peak instantaneous power (W) in a countermovement vertical jump with hands on hips for 26 male rugby players

Score	Test 1	Test 2
1	3443	3634
2	3880	3916
3	3927	3686
4	4195	4258
5	4845	4383
6	4292	4436
7	3150	3279
8	4364	4670
9	4215	4087
10	4324	4269
11	2042	4299
12	4177	3950
13	5324	5439
14	3423	3351
15	5216	5421
16	3753	4150
17	5270	5175
18	4021	3994
19	4633	4771
20	4959	4992
21	5003	5172
22	4502	4488
23	4607	4119
24	5560	5775
25	3708	3808
26	5458	5885

$r = 0.9255$ (based on the method shown in Table 2.3).
$s = 683.15$ = standard deviation of all of the test 1 and test 2 scores.
$SEM = s\sqrt{(1-r)} = 683.15(0.2729) = 186.43$ W.

peak instantaneous power scores in Table 2.4, $s = 683.15$ W and $r = 0.9255$. Therefore,

$$SEM = 683.15 \times \sqrt{(1-0.9255)} = 683.15 \times \sqrt{(0.0745)}$$
$$= 683.15 \times 0.2729 = 186.43 \text{ W}$$

TABLE 2.5 Guilford's classification of the magnitude of the coefficient of correlation (*r*) (Guilford 1950)

Magnitude of r	Degree of relationship
Less than 0.20	Slight, almost negligible
0.20–0.40	Low correlation, relationship definite but small
0.40–0.70	Moderate correlation, substantial relationship
0.70–0.90	High correlation, marked relationship
0.90–1.00	Very high correlation, very dependable relationship

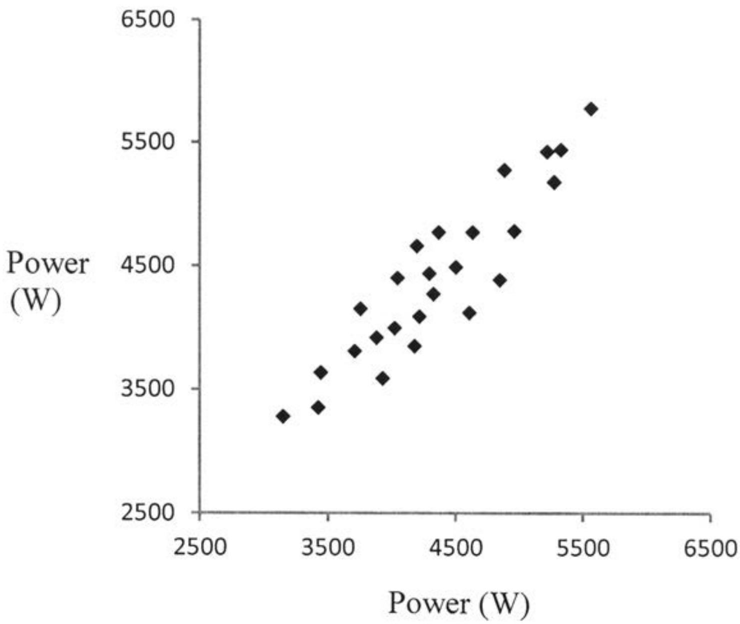


FIGURE 2.5 Scattergram of test-retest scores for peak instantaneous power (W) in a countermovement jump (hands on hips) for 26 male rugby players, 19–24 years of age.

The *SEM* is a standard deviation and in a normal distribution (Figure 2.3), 95% of the scores in the distribution lie within $\pm 1.96\sigma$ of the mean. Consequently, the tester can be 95% confident that when the test and retest data were collected the true score of any of the 26 players was within the range of test score $\pm$ 1.96 × 186.43 W, i.e. within the range of test score $\pm$ 365.40 W. The equation for the *SEM* indicates that the higher the reliability coefficient, the lower the *SEM* and, therefore, the greater the precision of measurement.

Validity

In addition to establishing the reliability and precision of measurement of a test, it is even more important to ensure that the test is valid, i.e. that it measures what it is supposed to measure. For example, if a student took an examination and all or most of the exam questions concerned topics that were not covered in the course, the student could complain that the exam did not measure what it was supposed to measure, i.e. that it was not valid (Safrit and Wood 1995). This is an example of a lack of content validity. There are four main types of validity: content, logical, construct and criterion (Thomas *et al.* 2005).

Content validity refers to the extent that the test measures content that is representative of the knowledge, skill or ability being assessed. For example, the extent that the content of an exam represents the content covered in the course determines the content validity of the exam. Similarly, the extent that the components of a test designed to assess the ability to play soccer reflect actual playing conditions determines the content validity of the test.

Logical validity, also known as face validity, refers to tests which clearly test the ability being tested. For example, measuring the ability to stand on one leg is clearly a test of static balance and measuring the ability to sprint 5 m from a standing start is clearly a test of speed of movement.

Construct validity refers to the extent that a test measures an ability or attitude that cannot be measured directly such as anxiety, intelligence, creativity and general athletic ability (Rampinini *et al.* 2007).

Criterion validity of a test refers to the extent that the results of the test correlate with the results of a criterion test that is known to be valid. There are two main types of criterion validity: predictive and concurrent. Predictive validity refers to the extent that the results of a test predict future criterion behaviour. For example, a test of intelligence might be used to predict scholastic achievement at some future date. Similarly, a soccer skills test might be used to predict soccer playing ability at some future date. Concurrent validity refers to the extent that the results of a test predict current performance in a criterion test. For example, mechanical power output is generally regarded as an important determinant of performance in all sports that require high rates of mechanical work, such as in sprinting, jumping and throwing (Owen *et al.* 2014). Consequently, athletes in many sports are regularly subjected to tests of power output within training programmes. There is general agreement that the most valid test of human mechanical power output is peak instantaneous power output in a countermovement vertical jump (Davies and Rennie 1968). However, administration of this test requires expensive equipment and trained personnel on a one-to-one basis. Consequently, the test is impractical for the majority of athletes. Not surprisingly, many attempts have been made to develop tests that have concurrent validity with the criterion test of Davies and Rennie (1968), but which require little or no equipment and are easy to administer in non-laboratory settings (Duncan *et al.* 2008; Quagliarella *et al.* 2011). These tests are referred to as 'field tests'.

In addition to field tests to assess power output, field tests have been developed to assess other fitness components. For example, the criterion method of measuring a person's per cent body fat involves weighing the subject underwater. This requires special laboratory facilities and trained personnel not normally available to most athletes. Consequently, several field tests based on skinfold measurements have been developed to predict criterion test scores for per cent body fat (Wilmore *et al.* 2008). Similarly, field tests based on bench stepping and distance run in a certain time have been developed to predict criterion test measures of cardiorespiratory function (McArdle *et al.* 2010). In Practical Worksheet 19, the concurrent validity of a standing long jump field test is determined with respect to the Davies and Rennie (1968) criterion test of power output.

> **BOX 2.7**
>
> There are four main types of validity: content, logical, construct and criterion.

The correlation between the scores in a field test and the corresponding criterion test scores is referred to as the 'validity coefficient' of the field test with respect to the criterion test. The higher the validity coefficient, the more linear the relationship between the two sets of scores, i.e. the more the scattergram approximates a straight line. In a perfect positive relationship (Figure 2.4a) and a perfect negative relationship (Figure 2.4b), all of the points in the corresponding scattergram lie on the same straight line. Consequently, each score on one variable is associated with a unique score on the other variable, i.e. a score on one variable can be used to predict a score on the other variable with 100% accuracy. For example, the first three columns of Table 2.6 show the time (t), speed (v) and distance (d_1) data corresponding to a man who starts from rest and walks at a constant speed of $v = 2$ m/s along a straight, level track for a period of 10 s. As $d_1 = v{\cdot}t$ and v is constant, there is a perfect positive correlation between distance and time. Consequently, the scattergram of distance-time data is a straight line as shown in Figure 2.6a. After walking for 10 s at 2 m/s the man will have travelled a distance of 20 m (2 m/s × 10 s = 20 m). If, on another occasion, the man walks on the same straight, level track at a constant speed of 2 m/s for 10 s from a start point that was 12 m forward of the original start point, his distance from the original start point at any particular point in time would be given by $d_2 = v{\cdot}t + 12$. Figure 2.6b shows the corresponding scattergram and the straight line joining the points. After walking for 10 s the man would be 32 m from the original start point:

$$d_2 = v{\cdot}t + 12 = (2 \text{ m/s} \times 10 \text{ s}) + 12 \text{ m} = 20 \text{ m} + 12 \text{ m} = 32 \text{ m}$$

TABLE 2.6 Distance–time data for a man who starts from rest and walks along a straight, level track at a constant speed of 2 m/s for 10 s. d_1, starting from a reference point where distance is zero; d_2, starting from a point that is 12 m forward of the reference point.

Time *(t)* (s)	Speed *(v)* (m/s)	Distance (d_1) (m)	Distance (d_2) (m)
0	2	0	12
1	2	2	14
2	2	4	16
3	2	6	18
4	2	8	20
5	2	10	22
6	2	12	24
7	2	14	26
8	2	16	28
9	2	18	30
10	2	20	32

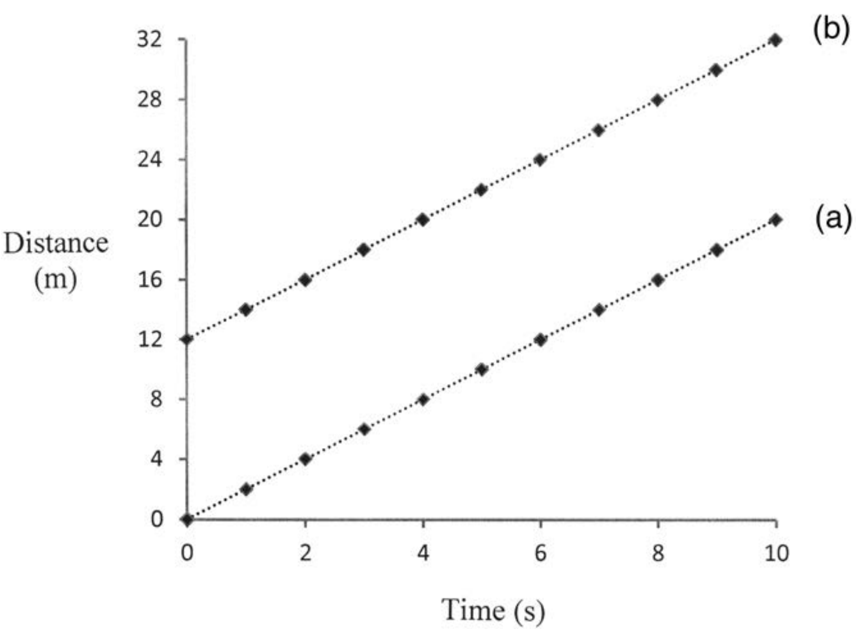


FIGURE 2.6 Distance–time graphs of a man who starts from rest and walks along a straight, level track at a constant speed of 2 m/s for 10 s. (a) Starting from a reference point where distance is zero. (b) Starting from a point that is 12 m forward of the reference point.

The equation of the line, $d_2 = v{\cdot}t + 12$, is a specific example of the general equation for a straight line relationship between two variables. The general equation is referred to as a linear function and is expressed as follows:

$$Y = a{\cdot}X + b$$

where

Y = the value that is to be calculated by the equation. Y is usually referred to as the dependent variable because its value depends upon the value of X.

X = the value that is known and is used to calculate Y. X is usually referred to as the independent variable.

a = a multiplicative constant that defines the rate at which Y changes as X changes, i.e. the gradient of the line.

b = an additive constant that gives a reference point from which changes in Y are measured. b is the value of Y when $X = 0$, i.e. the value of Y at the point where the line crosses (or intercepts) the Y axis. b is usually referred to as the intercept.

In the distance-time graph in Figure 2.6a, $b = 0$. Consequently the linear function shown in Figure 2.6a is $Y = 2X$, where Y represents distance in metres and X represents time in seconds; when $X = 0$, $Y = 0$. In the distance-time graph in Figure 2.6b, $b = 12$. Consequently the linear function shown in Figure 2.6b is $Y = 2X + 12$. In this case, when $X = 0$, $Y = 12$ m.

Regression

Unlike the perfect relationship ($r = 1$) between distance and time shown in Figure 2.6, validity coefficients between field and criterion test measures concerning physical fitness, anthropometric and behavioural variables are usually less than perfect. For example, the second and third columns of Table 2.7 show the vertical jump height scores (field test) and mechanical power output in a vertical jump scores (criterion test) for 25 young, physically active men. The corresponding scattergram (Figure 2.7) shows a clear trend for power to increase as jump height increases. This trend is reflected in the validity coefficient between the field and criterion test scores of $r = 0.7694$. The scattergram indicates that any particular jump height score is associated with a range of power scores. For example, a jump height of 0.45 m is associated with power scores in the approximate range of 3500 W to 4500 W. Consequently, based on this data, prediction of power from jump height would be subject to a margin of error. To define the margin of error, it is necessary to derive the linear function (equation of a straight line) that best represents the trend shown in the scattergram. In the absence of any other statistics, the best estimate of power for any particular jump height score would be the mean power score, i.e. 3955.82 W. This is shown as a dotted horizontal line in Figure 2.8 where $Y = 3955.82$ W.

If the relationship between the power scores and the jump height scores was perfectly linear, an increase in one standard deviation in jump height would be accompanied by an increase in one standard deviation in power. In this case, the gradient of the linear function would be equal to the standard deviation of the power scores divided by the standard deviation of the jump height scores and the linear function between power and jump height would be as follows:

$$Y = \overline{Y} + \frac{\sigma_Y}{\sigma_X}(X - \overline{X}) \qquad \text{Eq. 2.1}$$

where
Y = predicted power score
X = the jump score used to predict Y
$\overline{X}$ = the mean of the jump height scores = 0.439 m
$\overline{Y}$ = the mean of the power scores = 3955.82 W
σ_X = the standard deviation of the jump height scores = 0.053 m
σ_Y = the standard deviation of the power scores = 487.987 W

Consequently,

$$Y = 3955.82 + \frac{487.987}{0.053}(X - 0.439)$$
$$Y = 3955.82 + 9207.302\ (X - 0.439)$$
$$Y = 3955.82 + 9207.302X - 4042$$
$$Y = 9207.302X - 86.18 \qquad \text{Eq. 2.2}$$

The line corresponding to equation 2.2 is shown as the dotted oblique line in Figure 2.8. The scatter of the data points around this line is, in general, much lower than the scatter of the data points around the line Y = 3955.82 W, i.e. in general, the data points are much closer to the line Y = 9207.302 – 86.18 than they are to the line Y = 3955.82 W. Scatter is measured as the sum of the squared deviations of the actual Y scores from the corresponding predicted Y scores; the lower the sum of the squared deviations, the more accurate the predicted criterion test scores across the full range of possible field test scores. For each particular set of field test–criterion test data points, there is a linear function for which the sum of squared deviations is least. This line, which is called the least squares linear best fit, provides the most accurate predictions across the full range of possible field test scores. The linear function of the least squares linear best fit can be derived by adjusting the gradient of equation 2.1 by applying the validity coefficient as follows:

$$Y' = \overline{Y} + r\frac{(\sigma_Y)}{\sigma_X}(X - \overline{X}) \qquad \text{Eq. 2.3}$$

where
Y' = Y prime = predicted score from a non-perfect linear relationship
r = validity coefficient = 0.7694

Consequently,

$$Y' = 3955.82 + 0.7694\frac{(487.987)}{0.053}(X - 0.439)$$
$$Y' = 3955.82 + 0.7694(9207.302)(X - 0.439)$$
$$Y' = 3955.82 + 7084.098X - 3109.919$$
$$Y' = 7084.098X + 845.9 \qquad \text{Eq. 2.4}$$

The line corresponding to equation 2.4 is shown as the full oblique line in Figure 2.8. When the validity coefficient is less than 1.0, the gradient of the least squares linear best fit (Equation 2.4) is always depressed or regressed relative to that of the perfect linear function (Equation 2.2). For this reason, equation 2.4 is called a regression equation. When prediction is based on a single field variable, as in equation 2.4, the process is referred to as simple regression. When prediction is based on two or more field variables, the process is referred to as multiple regression. Multiple regression is beyond the scope of this book.

TABLE 2.7 Determination of the regression line equation and standard error of the estimate (*SEE*) for predicting peak instantaneous power in a countermovement jump with hands on hips (criterion test) from height jumped in a countermovement jump with hands on hips (field test). The subjects were twenty-five, young, physically active men.

Subject	Jump height *(X)* (m)	Power *(Y)* (W)	Predicted Power *(Y')* (W)	$d = Y - Y'$ (W)	d^2 (W^2)
1	0.352	3373.4	3339.50	33.90	1149.04
2	0.359	3126.1	3389.09	–262.99	69164.36
3	0.391	3276.9	3615.78	–338.88	114841.22
4	0.432	3439.9	3906.23	–466.33	217463.98
5	0.462	3655.5	4118.75	–463.25	214603.60
6	0.468	4216.9	4161.26	55.64	3096.05
7	0.504	4073.8	4416.28	–342.48	117296.24
8	0.435	4027.2	3927.48	99.71	9943.55
9	0.491	4731.5	4324.19	407.31	165899.71
10	0.395	3890.4	3644.12	246.28	60654.47
11	0.362	3549.6	3410.34	139.26	19392.38
12	0.442	3647.5	3977.07	–329.57	198617.25
13	0.414	4150.9	3778.72	372.18	138520.50
14	0.446	4378.9	4005.41	373.49	139496.49
15	0.476	4570.9	4217.93	352.97	124587.36
16	0.515	4315.2	4494.21	–179.01	32044.74
17	0.475	3845.9	4210.84	–364.95	133185.98
18	0.535	4981.6	4635.89	345.71	119513.72

(Continued)

TABLE 2.7 (Continued)

Subject	Jump height *(X)* (m)	Power *(Y)* (W)	Predicted Power *(Y')* (W)	$d = Y - Y'$ (W)	d^2 (W^2)
19	0.462	3655.5	4118.75	–463.25	214603.60
20	0.401	3464.5	3686.62	–222.12	49338.76
21	0.384	3588.4	3566.19	22.21	493.12
22	0.435	4027.2	3927.48	99.72	9943.55
23	0.521	4651.6	4536.71	114.88	13198.55
24	0.372	3890.4	3481.18	409.21	167457.36
25	0.455	4365.8	4069.16	296.64	87992.57
	$\overline{X} = 0.439$	$\overline{Y} = 3955.82$			$\sum d^2 =$ 2332498.20
	$\sigma_X = 0.053$	$\sigma_Y = 487.987$			

$r = 0.7694$

$$Y' = \overline{Y} + r\frac{(\sigma_Y)}{\sigma_X}(X - \overline{X}) = 7084.098X + 845.9$$

$$SEE = \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{2332498.20}{25}} = \sqrt{93299.93} = 305.45\,\text{W}$$

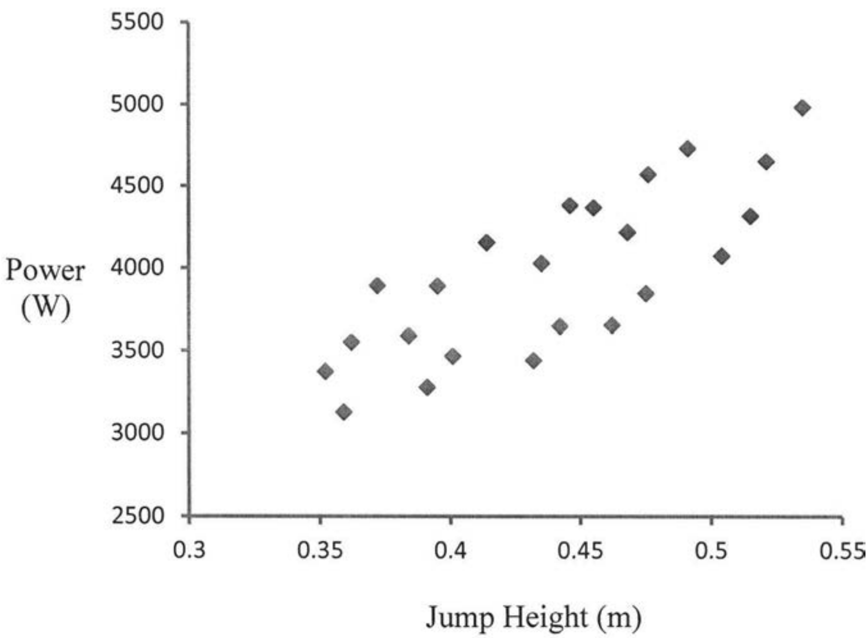


FIGURE 2.7 Scattergram of scores for vertical jump height (field test) and peak instantaneous power in a vertical jump (criterion test) for 25 physically active young men. $r = 0.7694$.

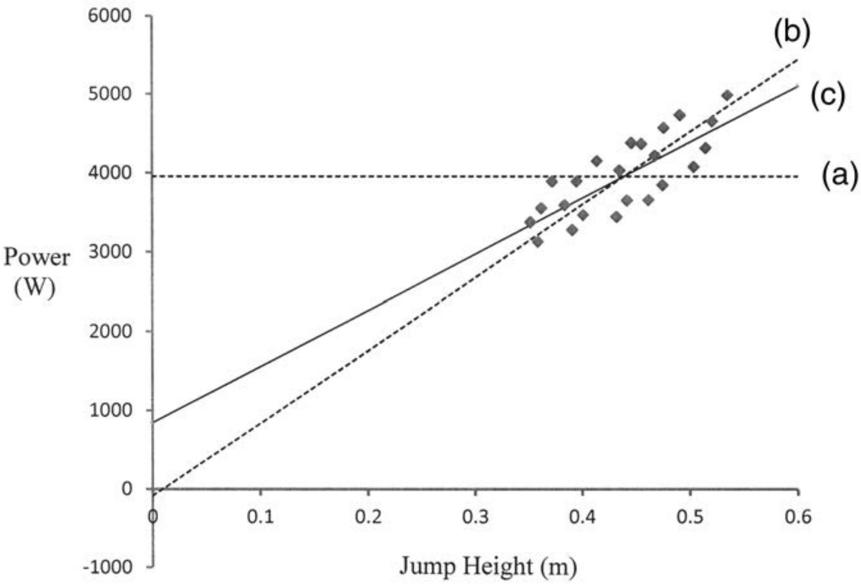


FIGURE 2.8 Simple regression. (a) The line $Y = 3955.82$. (b) The linear function $Y = 9207.302X - 86.18$ that would exist between power and jump height if the scores were perfectly related. (c) The linear function of the regression line between power and jump height, $Y = 7084.098X + 845.9$.

Standard error of the estimate

Just as the standard error of measurement provides an estimate of the precision of measurement of a test, a statistic called the standard error of the estimate (*SEE*) provides an estimate of the precision of prediction of a regression equation. The *SEE* is the standard deviation of the distribution of deviations of the actual *Y* scores from the corresponding predicted *Y* scores.

$$SEE = \sqrt{\frac{\sum d^2}{N}}$$

where
$\sum d^2$ = the sum of the squared deviations, i.e. the sum of $(Y - Y')^2$
N = the number of scores

The fourth column of Table 2.7 shows the predicted power scores based on equation 2.4. The fifth column of Table 2.7 shows the deviations of the predicted power scores from the actual power scores. The deviations are shown graphically in Figure 2.9. The sixth column of Table 2.7 shows the squares of the deviations together with the sum of the squared deviations.

$$SEE = \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{2332498.2}{25}} = \sqrt{93299.93} = 305.45 \text{ W}$$

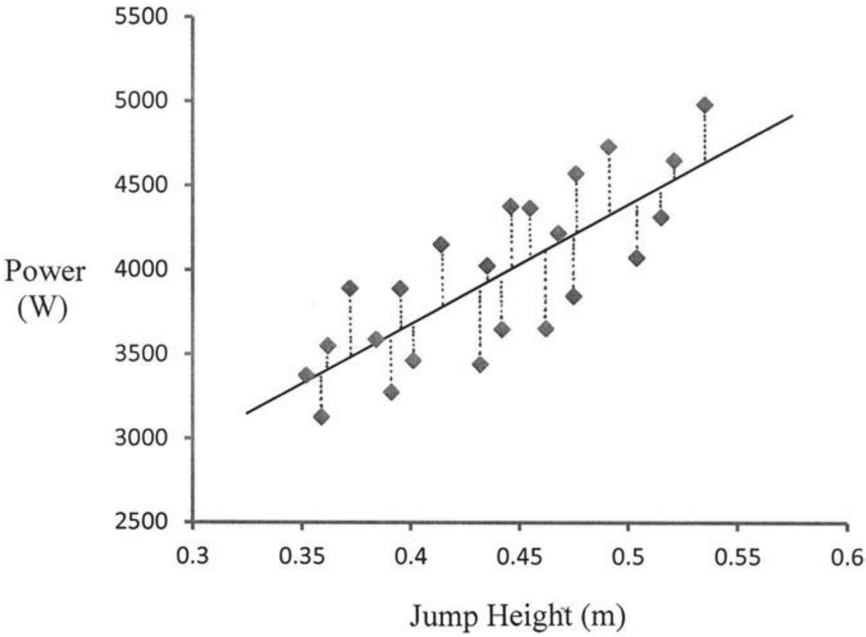


FIGURE 2.9 The deviations of the actual power scores from the corresponding predicted power scores are represented by the vertical distances between the actual power scores and the regression line. The regression line is the line through the scattergram for which the sum of the squares of the deviations is least.

In a normal distribution (Figure 2.3), 95% of the scores in the distribution lie within $\pm 1.96\sigma$ of the mean. Consequently, the tester can be 95% confident that when the field test and criterion test data were collected, the true criterion power score of any of the 25 participants was within the range of predicted power score $\pm$ 1.96 × 305.45 W, i.e. within the range of predicted power score $\pm$ 598.68 W.

> **BOX 2.8**
>
> Together with the standard error of the estimate, the regression line provides the most accurate estimate of a criterion test score based on a field test score.

References

Davies, C. T. M. and Rennie, R. (1968) 'Human power output', *Nature* 217:770–771.

Duncan, M. J., Lyons, M. and Nevill, A. M. (2008) 'Evaluation of peak power prediction equations in male basketball players', *Journal of Strength and Conditioning Research* 22(4):1379–1381.

Guilford, J. P. (1950) *Fundamental statistics in psychology and education*, 2nd edn. New York: McGraw-Hill Book.

McArdle, W. D., Katch, F. L. and Katch, V. L. (2010) *Exercise physiology: Energy, nutrition, and human performance*, 7th edn. Philadelphia: Lippincott Williams & Wilkins.

Owen, N., Watkins, J., Kilduff, L., Bevan, H. R. and Bennett, M. (2014) 'Development of a criterion method to determine peak mechanical power output in a countermovement jump', *Journal of Strength and Conditioning Research* 28(6):1552–1558.

Quagliarella, L., Sasanelli, N., Belgiovine, G., Moretti, L. and Moretti, B. (2011) 'Power output estimation in vertical jump performed by young male soccer players', *Journal of Strength and Conditioning Research* 25(6):1638–1646.

Rampinini, E., Bishop, D., Marcora, S., Ferrari-Bravo, D., Sassi, R. and Impellizzeri, F. M. (2007) 'Validity of simple field tests as indication of match-related physical performance in top-level professional soccer players', *International Journal of Sports Medicine* 28(2):228–235.

Safrit, M. J. and Wood, T. M. (1995) *Introduction to measurement in physical education and exercise science*, 3rd edn. St Louis, MO: Mosby-Year Book.

Spiegel, M. R., Schiller, J. J. and Srinavasan, R. A. (2013) *Probability and statistics*, 4th edn. McGraw-Hill: New York.

Stevens, S. S. (1946) 'On the theory of scales of measurement', *Science* 103(2684):677–680.

Thomas, J.R., Nelson, J.K. and Silverman, S. J. (2005) *Research methods in physical activity*, 5th edn. Champaign, IL: Human Kinetics.

Weather Online. Available: www.weatheronline.co.uk/

Wilmore, J. H., Costill, D. L. and Kenney, W. L. (2008) *Physiology of sport and exercise*, 4th edn. Champaign, IL: Human Kinetics.

Chapter 4

Regulation of anti-doping in sport

International and national operational frameworks

Neil Chester and Nick Wojek

4.1 Introduction

Ever since drug use has been recognised as a significant issue in sport the need to regulate it has been seen as an important step in safeguarding both the welfare of athletes and the integrity of sport. As the prevalence of doping grew it was evident that sports governing bodies needed to impose rules and regulations to provide clear direction regarding the use of drugs by athletes. Following on from this was a clear need to impose a framework by which governance and sanctions may be applied. Whilst the need for regulation was apparent, the difficulty in establishing such an operational framework was evident in light of the numerous sports governing bodies and cultural differences that exist. In recent years the creation of the World Anti-Doping Agency (WADA) has helped in providing much uniformity to the anti-doping movement. Nevertheless, challenges remain in ensuring parity from an anti-doping governance perspective both nationally and internationally across different sports.

4.2 Why regulate drug use in sport?

From Chapter 1 it is clear that there are a number of reasons why athletes might use drugs. It is also evident that there exists a distinction between those used to enhance performance and those used for what they have been designed to do, that is, treat illness or injury. It is for this distinction that anti-doping regulation exists, to deter athletes from using drugs to enhance performance. Whether regulation should exist at all is typically viewed as an ethical issue and has been the topic of much philosophical debate.

Maintenance of the proverbial 'level playing field' is often quoted as a major reason to prohibit the use of drugs in sport, as those who use performance-enhancing drugs are deemed to have an advantage over those who don't. However, it is argued that even without drugs the 'level playing field' does not exist due to both biological and environmental inequalities (Kayser et al., 2007). Debate often surrounds the fact that the use of many performance-enhancing drugs is not illegal under state law yet is prohibited under the rules laid down by WADA. This generally means that those not competing in organised sport may freely use such drugs to enhance image and performance, and not fear any sanctions. There is clearly an ethical issue surrounding the use of drugs for non-therapeutic purposes, but this is often left to an individual decision and one that is often not legislated for, outside of organised sport. This is unless such drugs are deemed to have a significant impact on public health, for example some recreational psychoactive drugs such as amphetamine or cocaine.

A major argument in support of the regulation of drug use in sport is related to the welfare of athletes in terms of protecting their health. Despite the lack of empirical evidence, it is widely accepted that athletes who use drugs (and methods) to enhance sports performance are putting their health at significant risk. However, the opposing argument might be that participation in sport, particularly at an elite level, increases the chances of developing serious health problems such as injury (Kayser et al., 2007). One might argue that the risks of participating in sport might be made more apparent, allowing an individual to make a more informed decision as to whether they partake or not.

Whether one agrees with the ethical arguments put forward to justify the legislation of drug use in organised sport, the bottom line remains that the use of performance-enhancing drugs, and in most cases recreational drugs, is against the rules of sport. Therefore, much like the handball rule in soccer, the use of such drugs is prohibited. The whole basis of competitive sport relies on rules and without such rules sport would cease to function. The ethical arguments around drug use in sport therefore become secondary but nonetheless no less important.

4.3 The history of the anti-doping movement

Anti-doping has been a reactive movement in response to key incidents in sport that have highlighted not only the use of drugs as performance enhancers but also the dangers surrounding their use from a health perspective.

The International Association of Athletics Federations (IAAF), formerly known as the International Amateur Athletics Federation, was the first sports federation to implement anti-doping regulations when it banned stimulants in 1928. However, it wasn't until the 1960s following the untimely death of the Danish cyclist, Knud Jenson in the Rome Olympics, allegedly as a consequence of amphetamine use, that other sports federations took significant steps to legislate against the use of drugs in sport. In 1966 the Fédération Internationale de Football Association (FIFA) introduced a list of prohibited substances and the following year the Union Cycliste Internationale (UCI) and the International Union of Modern Pentathlon (UIPM) followed suit (Mazzoni et al., 2011). In the same year the IOC formalised its battle against drug misuse in sport by establishing a Medical Commission to oversee doping matters and introduce anti-doping regulations.

In 1967 saw the death of Tom Simpson, the British cyclist who died shortly after his collapse close to the summit of Mont Ventoux during the thirteenth stage of the Tour de France, which was attributed to the use of amphetamine and alcohol. At this time the UCI and FIFA were the first to introduce drug tests as a deterrent to their athletes in their respective World Championships in 1966. The IOC followed suit in 1968 by introducing a list of prohibited substances and drug testing in time for the Winter Olympic Games in Grenoble and the Summer Games in Mexico City. Initial urine tests could only detect the use of stimulants such as amphetamine and it was not until the mid-1970s that a test was established for the detection of anabolic androgenic steroids (AAS). Further advances in analytical chemistry enabled a growing list of prohibited substances to be detected in urine, however problems still remained in terms of the effectiveness of doping control methods.

In 1988, one of the most famous episodes of drug misuse in sport led the anti-doping movement to consider its efficacy once more. The positive test, for the AAS stanozolol by the Canadian track and field athlete, Ben Johnson, immediately after the

100m Olympic final, led to a large review of the use of performance-enhancing drugs by athletes. This review was led by the Canadian lawyer Charles Dubin and was known as the Dubin Inquiry. The inquiry was to last one year and unearthed widespread doping amongst athletes as well as marked inadequacies from a doping control perspective. It also put forward numerous recommendations in an attempt to control performance-enhancing drug use in sport (Moriarty et al., 1992).

In addition to improved testing and stricter penalties to act as a deterrent to those partaking in doping or considering it in the future, the Dubin Inquiry also put forward several ambitious recommendations that would attempt to stem the tide of widespread doping. Such recommendations focused on maintaining ethical standards including changing the emphasis away from extrinsic rewards, such as gold medals, towards intrinsic rewards and incorporating ethics and morality into coach education (Moriarty et al., 1992). By and large many of the recommendations made in the report form the basis of the anti-doping education that we see today. However, as elite sport attracts a wider audience the rewards that come with success continue to grow and therefore it would appear that a huge cultural change would be needed for motivation to shift towards those that are more intrinsically based.

Although the failed drug test of Ben Johnson put doping in sport in the media spotlight this was to be no isolated case. Indeed, less than two years later reports of systematic doping in the former German Democratic Republic (GDR) were to be uncovered as reunification of Germany was reached. Shocking reports were uncovered that provided evidence of a state-run doping programme that would be central to the GDR success from the mid-1960s until reunification of Germany in 1990.

Despite the horrific reports of systematic doping in the GDR it is interesting to note that the sporting community did relatively little in the aftermath to address such atrocities. The international federations and the IOC did not address the humanistic issues that the investigation into the GDR-system unearthed nor did they look to rescind any of the medals or records that were achieved by known doped athletes. Franke and Berondonk (1997) provide a comprehensive review of the doping practices in the former GDR and in doing so offer a sobering reflection that doping was unlikely to have been isolated to such a small state or that other individuals with similar support would not have done the same.

It was not until after the events of the Tour de France in 1998, where large-scale team doping was uncovered, that the anti-doping movement would be pushed to make a monumental shift and establish the World Anti-Doping Agency (WADA). This allowed anti-doping as a movement to function both independently of sports federations and, most importantly, globally to harmonise the fight against doping in sport. On 10 November 1999, the IOC convened the first World Conference on Doping in Sport, held in Lausanne, Switzerland, which led to the formation of WADA.

Throughout the proceeding years WADA, following extensive consultation, produced the first draft of the World Anti-Doping Code (WADC) which would provide the framework for the anti-doping movement throughout the twenty-first century. The second World Conference on Doping in Sport was held in November 2003 in Copenhagen. In addition to major international sports federations, stakeholders from 80 governments were represented, who formally agreed the Copenhagen Declaration on Anti-Doping in Sport. This declaration was a formal acceptance of the WADC (and of WADA) by governments which was to come into effect on 1 January 2004. As part of the Code, international standards including a new List of Prohibited Substances and Methods (formally produced by the IOC Medical Commission) were introduced.

Since the WADC is a non-government document and therefore not legally binding for governments, the Copenhagen Declaration was further developed by the United Nations Educational, Scientific and Cultural Organisation (UNESCO). Indeed, the UNESCO International Convention against Doping in Sport was developed to provide an internationally recognised legal framework for governments to attend to doping in sport and to recognise the WADC. On 19 October 2005, the UNESCO Convention was adopted and took effect from 1 February 2007.

Further World Conferences on Doping in Sport in 2007 and 2013 have seen two revisions of the WADC and International Standards to adapt to the ever-changing doping landscape and further harmonise anti-doping regulations.

4.4 Anti-doping structure

Anti-doping from an organisational perspective is led by WADA, which links both international sporting organisations and state governments. There are few institutions that successfully combine such varied organisations with a unified goal to protect athletes against doping and provide a level playing field globally, across all sports. Clearly by combining both sports organisations and governments the fight against doping in sport benefits from not only a unified and consistent approach, but also from the resources that each organisation individually brings to the table.

Whilst WADA brings together both sports organisations, such as the IOC, and state governments in a unique partnership, its role is largely to manage the World Anti-Doping Program which has been developed to ensure that the WADC is adhered to by all those who sign up to it (Figure 4.1). As an independent body, WADA facilitates and monitors the efforts of all signatories (i.e. governments and sports federations) in terms of their compliance with the WADC and, where necessary, sanctions maybe imposed where non-compliance is evident. A set of international standards have also been developed to help to operationalise the Code together with models of best practice which provide guidelines and solutions for many anti-doping issues.

Regional and national anti-doping organisations (RADOs and NADOs) are in place to ensure that countries comply with the WADC. States that do not have the funds to commit to a fully operational NADO may be served by a RADO which ensures Code compliance across several states within a particular region. In the UK, UK Anti-Doping is the NADO responsible for all anti-doping matters and ensures effective governance of a national anti-doping programme set up to enable compliance with the Code.

The power afforded to WADA lies in the fact that the participation of specific sports and nations at major international competitions including the Olympics, Paralympics and World Championships (of various sports) is dependent on both the acceptance and implementation of the WADC.

From a legislative perspective, some countries have specific laws to help govern drug use in sport (e.g. France, Germany and Italy) whilst others have laws to help govern societal drug use as a whole and may be applied to doping (e.g. UK and USA). Nevertheless, the UNESCO convention is in place to provide a legal framework for those governments that have ratified it.

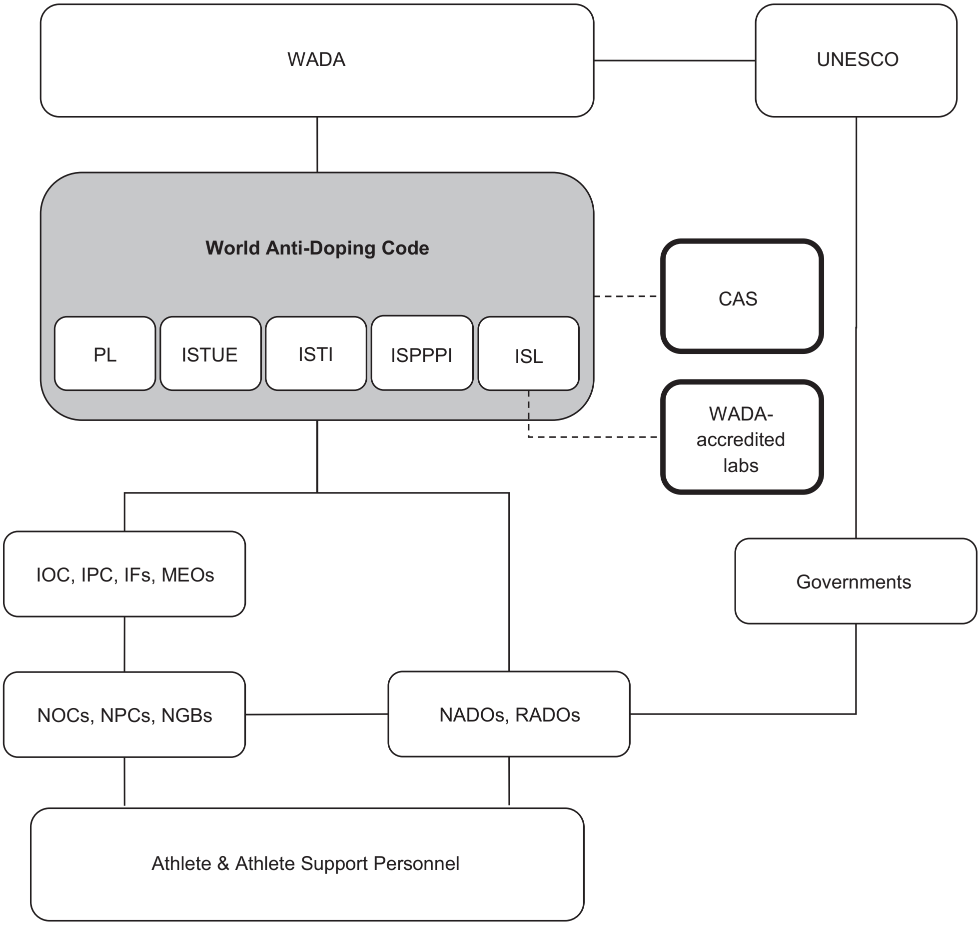


Figure 4.1 A schematic representation of the structure of anti-doping within the context of the World Anti-Doping Code

CAS, Court of Arbitration in Sport; IOC, International Olympic Committee; IPC, International Paralympic Committee; ISL, International Standard for Laboratories; ISPPPI, International Standard for Protection of Privacy and Personal Information; ISTUE, International Standard for Therapeutic Use Exemptions; ISTI, International Standard for Testing and Investigations; MEOs, Major Event Organisers; NADO, National Anti-Doping Organisation; NGBs, National Governing Bodies; RADO, Regional Anti-Doping Organisation; NOCs, National Olympic Committees; NPCs, National Paralympic Committees; PL, Prohibited List; UNESCO, United Nations Educational, Scientific and Cultural Organisation.

4.5 The World Anti-Doping Agency

In 1999 at the first World Conference on Doping in Sport, convened by the IOC in Lausanne, the World Anti-Doping Agency was founded and according to its constitution (WADA, 2016a), its purpose at an international level is to:

1 Promote and coordinate anti-doping in sport both in- and out-of-competition;
2 Reinforce ethical principles to underpin doping-free sport and protect the health of athletes;
3 Establish and update annually a list of prohibited substances and methods in sport;
4 Support and coordinate an out-of-competition drug testing programme;

5 Develop and harmonise a scientific approach to drug testing through technical standards and procedures in sampling and analysis;
6 Promote harmonised rules, disciplinary procedures and sanctions to combat doping in sport;
7 Develop an education programme to promote doping-free sport based on ethical principles; and
8 Promote and coordinate research into anti-doping.

Members of the WADA Foundation Board are made up, essentially, of representatives from the IOC and public authorities (i.e. government) to form an equal partnership. The president of WADA is an honorary position that lasts for a maximum of two 3-year terms. The position, appointed by the Foundation Board, alternates between representation from the IOC and public authorities. In addition to the Foundation Board an Executive Committee takes charge of the actual management and running of the organisation. There are also several additional committees with specialist roles, including: the Athlete Committee; the Education Committee; the Finance and Administration Committee; and the Health, Medical and Research Committee. Further groups have been formed to provide expert opinion in specialist areas and serve important roles with respect to the Code, including: The Prohibited List; Therapeutic Use Exemptions; Laboratories; Technical Document for Sport Specific Analysis; Gene Doping; and Ethical Issues.

As a partnership between governments and the IOC, WADAs budget is funded equally by both parties. In 2015 the WADA budget amounted to a total of over 26 million US dollars (WADA, 2016b).

4.6 The UNESCO International Convention against Doping in Sport

UNESCO introduced the International Convention against Doping in Sport on 1 February 2007. The convention provides the legal framework to enable governments to address anti-doping in sport. Whilst sports organisations may be able to progress so far with regards to anti-doping and sanctioning athletes there is a necessity for government to be able to support this work. Indeed, many doping programmes are so extensive that only governments have the authority to deal with them. Clearly to tackle a doping culture there is a need to focus on issues surrounding drug availability and distribution and to address the part in which athlete support personnel play in a doping case. Essentially the Convention highlights the importance of anti-doping and ensures that governments play a concerted effort in tackling the problem. According to the 2015 WADA Annual Report, 182 of 195 UNESCO member states, including all of the Americas and Europe, have ratified the UNESCO convention (WADA, 2016b).

4.7 The World Anti-Doping Code

Since its inception in 2004 over 660 organisations have signed up to the World Anti-Doping Code (WADA, 2017a) including, amongst others, the IOC, the International Paralympic Committee (IPC), International Sports Federations (IFs), National Olympic and Paralympic Committees (NOCs and NPCs), Regional and National Anti-Doping Organisations (RADOs and NADOs) and event organisations (e.g. Commonwealth Games Federation). The Code provides a universal standard for anti-doping practice which includes a wide range of activities ranging from drug testing to education and research. In addition to the WADC there are a number of mandatory documents which outline the international standards of operation for the key activities of anti-doping organisations and personnel, including:

1 The List of Prohibited Substances and Methods;
2 Therapeutic Use Exemptions;
3 Testing and investigations;
4 Laboratories;
5 Protection of privacy and personal information.

Models of best practice and guidelines relating to the Code and its implementation are also available to use by signatories, but are not mandatory (WADA, 2017b).

As the anti-doping landscape changes WADA must adapt accordingly and it does this most notably through regular revisions to the Code. Whilst the first WADC was introduced in 2004, the most recent revision of the Code was approved in 2013 at the fourth World Conference on Doping in Sport, in Johannesburg, following a widespread consultation process involving stakeholders.

The revised Code came into effect on 1 January 2015 and includes several changes to the 2009 version. The most significant changes include increased sanctions for those who commit a first doping offence, a recognition of the role of non-analytical evidence in investigations into potential ADRVs, and the role of athlete support personnel in doping offences. Further detail summarising the significant changes between the 2009 and 2015 Codes can be found in an overview document on the WADA website (WADA, 2013a).

Compliance with the Code is essential and thus a major element of the Code itself. Clearly in order for the anti-doping movement to function all signatories of the Code must adhere to the rules and regulations set out in the document. In doing so signatories must put in place policy that ensures that the rules, regulations and procedures set out in the Code are followed by all stakeholders. Governance relating to Code compliance is therefore an important aspect of WADA's role.

The following sections will include a brief description of the key elements of the WADC, including: doping control, education and research, and roles, responsibilities and compliance.

Doping control

A major portion of the Code is focused on anti-doping rules and regulations and the procedures required in order to enforce such rules. The Code outlines a wide range of rules and regulations that athletes must follow and sports organisations must implement. Further detail in relation to doping control procedure is provided in Chapter 5.

What is doping?

According to the Code (WADA, 2015a), doping is the occurrence of one or more of the following anti-doping rule violations (ADRVs):

1 The presence of a prohibited substance or its metabolites or markers in an athlete's sample;
2 The use of or attempted use by an athlete of a prohibited substance or method;
3 Refusing or failing to provide a sample, after notification without compelling justification;
4 Violation of applicable requirements regarding athlete availability for out of competition testing, including failure to file sufficient whereabouts information and missed tests;
5 Tampering or attempted tampering with any part of the doping control procedures;
6 Possession of prohibited substances or methods;
7 Trafficking or attempted trafficking in any prohibited substance or method;
8 Administration, attempted administration or assisting in the administration of any prohibited substance or method;
9 Intentional complicity (e.g. aiding, abetting, conspiring, covering up) to commit an ADR; and
10 Association in a professional or sport-related capacity of an athlete with support personnel who are serving a period of ineligibility or who have been convicted in a criminal proceeding for conduct that would constitute doping.

Proof of doping can be rather difficult to establish since many ADRVs do not involve the determination of a prohibited substance or use of a prohibited method via a positive drugs test. In attempts to establish non-analytical ADRV, evidence may be gathered from a variety of sources, including the admission by an athlete, the testimony by a third person as well as other documentary evidence. Where an ADRV is alleged to have occurred, an anti-doping organisation must establish that the proof is greater than the balance of probability for a case to proceed. The individual accused of an ADRV may then refute the claims at a hearing.

The Prohibited List

Whilst not all ADRVs involve the use of a prohibited substance or method the Prohibited List is fundamental to doping control since it clearly outlines what is deemed to be unacceptable both in- and out-of-competition (refer to Table 1.1 of this book). As with the Code, the Prohibited List is an evolving document. New pharmacological agents and potential performance-enhancing methods are continually being developed and becoming available and therefore the List must be updated regularly to keep pace with the rapidly changing landscape.

A specific expert committee has been established by WADA to oversee the development of the List and consider the inclusion or exclusion of specific substances or methods on an annual basis. Each year the List Expert Group initiates a consultation process to all WADA stakeholders asking them to consider potential modifications to the List. Modifications may take the form of changes to the terminology of the document (including the use of particular drug names) to help in the understanding and thus compliance of the List or changes to the actual content in terms of introducing or removing substances or methods to the List. In addition, the List Expert Group will also consider the inclusion of particular thresholds for specific substances or the reclassification of particular substances.

The criteria by which the List Expert Group make decisions as to whether a substance or method is to be considered for inclusion on the List are as follows:

1 The use of a substance or method has the potential to enhance performance in sport;
2 The use of a substance or method has the potential to adversely affect health; and
3 The use of a substance or method contravenes the 'spirit of sport'.

A substance or method is considered for inclusion by the List Expert Group if it meets any two of the above criteria. In addition, substances or methods that have the potential to mask the presence or use of a prohibited substance or method would also be considered for inclusion on the Prohibited List.

The List Expert Group will consider the comments made during the consultation process and also consider data made available to them from WADA-accredited laboratories, the WADA Monitoring Program and from research publications, particularly from those funded by WADA research grants.

The Monitoring Program

In addition to the List of Prohibited Substances and Methods, WADA also conducts a Monitoring Program which allows for monitoring of substances beyond that of the Prohibited List to determine patterns of use that may reflect possible misuse (WADA, 2016c). The programme typically includes substances that are not prohibited but have the potential to be misused in sport for possible performance-enhancing purposes. There are however a number of substances that are part of the Monitoring Program which are also present on the Prohibited List, although subject to only partial prohibition (i.e. only in-competition and/or above a particular threshold).

Therapeutic Use Exemptions

Clearly doping by athletes is not to be confused with the use of drugs for therapeutic purposes. Whilst drug use in sport is synonymous with doping there are numerous instances when drug use by athletes is entirely legitimate. Indeed, WADA have outlined a process whereby athletes may apply for, and be granted an exemption for the use of a particular prohibited substance or method specifically for therapeutic purposes. This process is known as Therapeutic Use Exemption (TUE) and is outlined in detail in the International Standard for Therapeutic Use Exemptions (WADA, 2015b). The procedure involved in obtaining a TUE requires an athlete to make a request with supporting medical evidence. Each application is then considered according to the guidelines set out by the International Standards document, by an independent TUE Committee that is formed by the respective NADO, IF or major event organiser. TUE Committees are made up of physicians with experience in the field of sports medicine and specialists with expertise in treating specific medical conditions. TUE Committees make decisions as to whether to grant a TUE based on the following four criteria (WADA, 2015b):

- The athlete would experience significant health problems without taking the prohibited substance or method;
- The therapeutic use of the substance would not produce significant enhancement of performance;
- There is no reasonable therapeutic alternative to the use of the otherwise prohibited substance or method; and
- The necessity for the use of the prohibited substance or method is not a consequence of the prior use of a substance or method which was prohibited at the time of such use.

Decisions made by each committee are then reported to WADA who may then decide to confirm or rescind the initial decision according to whether or not the TUE Committee complied with procedures set out in the International Standards document. There is also an appeals process which can be used by athletes to contest a Committee decision.

Testing

In order to enforce the rules set out in the Code and International Standards a robust testing procedure is necessary. Indeed, so that the Prohibited List may serve as an effective deterrent testing is required to be a covert operation that follows strict guidelines in terms of sampling and analysis as outlined in the International Standard for Testing and Investigations (WADA, 2016d) and for Laboratories (WADA, 2016e).

Sample analysis is carried out in specific laboratories that follow approved procedures and are formally accredited by WADA. Strict accreditation and regular assessment procedures ensure both validity and harmonisation of test results across all laboratories and that the highest standards are maintained. In recent times, a few laboratories have had their accreditations suspended and even revoked following non-conformities with the International Standard for Laboratories.

There are currently 34 WADA-accredited laboratories that are distributed in cities across the globe (WADA, 2017c). More than half (19 out of 34) of the accredited laboratories are based in Europe whereas the African and South American regions are considered underserved. As a result, WADA are discouraging the set-up of further laboratories in Europe and are exploring opportunities to establish additional laboratories in South America and Africa (WADA, 2013b).

Results management

Allied to the high standards maintained by a laboratory is a results management system which ensures a clear chain of events whereby an athlete is notified as soon as their sample returns an adverse analytical finding that is not supported by a valid TUE nor is there any reported deviation from the procedures as outlined in the International Standards for Testing or Laboratories. The athlete may then request the analysis of their B-sample which, if confirms the analytical results of the A sample an adverse analytical finding is established.

In the event of an adverse analytical finding there is a requirement for an anti-doping organisation to provide a fair hearing process to establish if an ADRV has occurred and what sanctions are to be imposed. Where the hearing confirms the evidence put forward by the anti-doping organisation and a sanction is applied an appeals process must be made available. A disciplinary case for doping might be dealt with, initially between the NADO, IF and athlete and a subsequent appeal might be lodged with the Court of Arbitration for Sport (CAS).

Sanctions

The Code clearly outlines the sanctions to be imposed on individuals, teams and sporting organisations implicated in an ADRV. The provision of clear sanctions is imperative to ensure harmonisation and act as a real deterrent to all those tempted by doping. Current sanctions include disqualification and forfeiture of points, medals and prizes where an ADRV has occurred in-competition. In addition, a period of ineligibility is typically imposed which varies according to the nature of the ADRV and whether the athlete has committed a previous ADRV. The criteria for increasing or reducing the standard four-year period of ineligibility can be rather complex and are outlined in Chapter 5. A public disclosure of those individuals committing an ADRV is arguably the most impactful from both a professional and personal perspective. An athlete who is deemed to have committed an ADRV should be publicly identified after such a time that a hearing and subsequent appeal may have been concluded.

In addition to sanctions imposed on individuals there are circumstances whereby a team may be sanctioned particularly in circumstances where two or more members of a team commit an ADRV. In such instances likely sanctions would include loss of points or disqualification from a competition.

Whereabouts

As a means of conducting an effective and efficient out-of-competition testing programme a system whereby athletes must report their prospective location for one hour a day, each day throughout their competitive career has been established. This system, where athletes notify either their IF or NADO of their location is known as 'whereabouts'. Only elite athletes competing at the highest level who are identified by their respective IF or NADO are required to undertake such reporting and as such are part of a 'registered testing pool'. The accuracy of the information provided is critical and those athletes who fail to file information or file inaccurate information (i.e. they are found not to be at the location, at the time that they have stated in their whereabouts information) on three separate occasions over a 12-month period will be deemed to have committed an ADRV.

Doping control for animals competing in sport

Sports that involve animals such as horse racing implement anti-doping rules and procedures specific to the animals via their IF. However, testing of humans is under the jurisdiction of WADA where the specific federation has signed up to the Code.

Statute of limitations

There is a period of ten years during which an alleged anti-doping violation may have taken place and throughout which action against an athlete may proceed. This duration of time acknowledges that sophisticated doping can sometimes take a long time to uncover. In addition, where multiple ADRVs are to be considered in determining the sanction (i.e. length of period of ineligibility) these must occur within the same ten-year period.

Education and research

A key element of the WADC is seen to be the promotion of proactive or preventative measures in an attempt to limit both intentional and unintentional ADRVs. In the Code both research and education are highlighted as key preventative measures in which all signatories are expected to engage.

WADA have established two funding streams in an attempt to promote research in the physiological and analytical sciences and the social sciences. Funding is assigned to the two programmes on an annual basis and researchers are encouraged to apply for funds through an application process involving the submission of a detailed project proposal. Following review by the appropriate WADA committee, approval is given to those projects that are deemed to answer important questions or add key information to new and emerging areas within the field of anti-doping. Funds committed to research since 2001 has been over 69 million US dollars with 8.8 million US dollars paid out in 2014 and 2015 (WADA, 2016b).

The criticism levelled at many research studies that attempt to examine the reputed effects of prohibited substances and methods is that findings are those from recreational or non-elite athletic subjects and thus not representative of elite athletes. Whilst there might be a drive to redress this issue there is a need to be cognisant of the issues that this might raise. Indeed recruiting athletes to studies that include the supplementation of a prohibited substance or use of a prohibited method should be avoided due to the potential performance-enhancing effects and adverse side-effects and also the potential of a failed drugs test as a consequence of participation in a research study (Howman, 2013).

Roles, responsibilities and compliance

Clearly one of WADA's key roles is to monitor and ensure that its signatories adhere to the rules and procedures set out in the WADC. Nevertheless, ultimately each organisation that is governed by the Code must not only ensure its own compliance with the Code but also that of those organisations that it governs. The consequences of non-compliance should also be visible, with a clear penalty for perpetrators such as ineligibility to participate in major competitions or indeed to bid to host such events (WADA, 2015a). There is a clear chain of command in terms of roles and responsibilities from a Code compliance perspective which filters down from WADA, the IOC and the IPC to IFs, NADOs, NOAs and NPAs and ultimately to NGBs and athletes. The chain of command which also includes Government and major event organisers, is complex (see Figure 4.1), yet essential if anti-doping is to be effective. Devolution of governance ensures that there is engagement with anti-doping at all levels.

Whilst acceptance of the Code by signatories has been particularly encouraging, Houlihan (2013) argues that the degree of compliance is less so. Clearly such an issue places the core WADA principle of harmonisation in doubt. Further efforts are required to address issues concerning the capacity and commitment of stakeholders in terms of Code compliance. Indeed, a Code compliance framework culminating in the drafting of a new International Standard on the Code Compliance of Signatories is currently being developed by WADA for 2018. This new International Standard will outline (WADA, 2017d):

- Code signatories' rights and responsibilities;
- The ways WADA supports signatories in achieving and maintaining Code compliance;
- A range of consequences that could be imposed in situations of non-compliance; and
- A process whereby the consequences can be imposed by an independent tribunal.

WADA has also recently introduced a Compliance Monitoring Program to strengthen its ability to monitor and scrutinise signatories' enforcement of the Code. Monitoring will be achieved through questionnaires sent to signatories, the independent collection of information that is available to WADA (e.g. through investigations, tip-offs, and the mandatory input of drug test results into WADA's Anti-Doping Administration and Management System by Anti-Doping Organisations), and through audits of NADOs' and IFs' anti-doping programmes.

4.8 Court of Arbitration for Sport

CAS has an important role within the structure of anti-doping since it provides an independent means, where necessary, for the resolution of disputes associated with ADRVs and related sanctions. It is available to both athletes and federations for all sports-related disputes including those linked to disciplinary charges such as doping and those that are commercial in nature such as sponsorship and relationship disputes between athletes, coaches, clubs and agents.

The concept of CAS was introduced by the IOC and established in 1984. Based in Lausanne it was comprised of 60 members that were appointed by the IOC, IFs and national Olympic Associations. However, the independence of CAS was later questioned and in 1994, following major reform, it became both structurally and financially autonomous.

4.9 Summary

The anti-doping movement has made huge strides over recent decades in an attempt to address the increasingly complex nature of doping practice. The establishment of policies and practices to address doping in sport is now led by the World Anti-Doping Agency through the World Anti-Doping Code. International Federations, international and national Olympic and Paralympic committees, major event organisations and national anti-doping organisations are all signatories to the Code. Governments accept the Code through signing up to the UNESCO Convention which provides the legal framework to enable governments to address anti-doping in sport. Whilst sports organisations may be able to progress so far with regards to anti-doping and sanctioning athletes there is a necessity for governments to be able to support this work aligning their domestic policies with the Code. However, challenges exist to guarantee that such a model can be implemented across all states thus ensuring harmonisation, a fundamental objective of the Code.

4.10 References

Franke, W.W. and Berendonk, B. (1997) Hormonal doping and androgenization of athletes: A secret program of the German Democratic Republic government. *Clinical Chemistry* 43: 1,262–1,279.

Houlihan, B. (2013) Achieving compliance in international anti-doping policy: An analysis of the 2009 World Anti-Doping Code. *Sport Management Review* 17(3): 265–276.

Howman, D. (2013) Scientific research using elite athletes: WADA point of view. *Journal of Applied Physiology* 114(10): 1,365.

Kayser, B., Mauren, A. and Miah, A. (2007) Current anti-doping policy: a critical appraisal. *BMC Medical Ethics* 8(2): 1–10.

Moriarty, D., Fairall, D. and Galasso, P.J. (1992) The Canadian Commission of Inquiry into the use of drugs and banned practices intended to increase athletic performance. *Journal of Legal Aspects of Sport* 2(1): 23–31.

WADA (2013a) Significant changes between the 2009 Code and 2015 Code. Available at: https://www.wada-ama.org/en/resources/the-code/significant-changes-between-the-2009-code-and-the-2015-code (Accessed on 28 July 2017).

WADA (2013b) Laboratory Network Strategy. Available at: https://www.wada-ama.org/en/resources/science-medicine/laboratory-network-strategy (Accessed on 28 July 2017).

WADA (2015a) World Anti-Doping Code 2015. Available at: https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code (Accessed on 28 July 2017).

WADA (2015b) International Standard for Therapeutic Use Exemptions. Available at: https://www.wada-ama.org/en/resources/therapeutic-use-exemption-tue/international-standard-for-therapeutic-use-exemptions-istue (Accessed on 28 July 2017).

WADA (2016a) Constitutive Instrument of Foundation of the World Anti-Doping Code. Available at: https://www.wada-ama.org/en/resources/legal/revised-statutes (Accessed on 28 July 2017).

WADA (2016b) 2015 Annual Report. Available at: https://www.wada-ama.org/en/resources/finance/annual-report (Accessed on 28 July 2017).

WADA (2016c) The 2017 Monitoring Program. Available at: https://www.wada-ama.org/en/resources/science-medicine/monitoring-program (Accessed on 28 July 2017).

WADA (2016d) International Standard for Testing. Available at: https://www.wada-ama.org/en/resources/world-anti-doping-program/international-standard-for-testing-and-investigations-isti-0 (Accessed on 28 July 2017).

WADA (2016e) International Standard for Laboratories. Available at: https://www.wada-ama.org/en/resources/laboratories/international-standard-for-laboratories-isl (Accessed on 28 July 2017).

WADA (2017a) Code Compliance. Available at: https://www.wada-ama.org/en/what-we-do/the-code (Accessed on 28 July 2017).

WADA (2017b) Model Rules, Guidelines and Protocols. Available at: https://www.wada-ama.org/en/model-rules-guidelines-and-protocols (Accessed on 28 July 2017).

WADA (2017c) List of WADA Accredited Laboratories. Available at: https://www.wada-ama.org/en/resources/laboratories/list-of-wada-accredited-laboratories (Accessed on 28 July 2017).

WADA (2017d) Code compliance and development of an International Standard for Code Compliance by Signatories. Available at: https://www.wada-ama.org/en/resources/the-code/back grounder-code-and-is-review (Accessed on 28 July 2017).