

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,000

Open access books available

125,000

International authors and editors

140M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



## Prediction of Sports Injuries by Mathematical Models

Juan Carlos de la Cruz-Márquez, Adrián de la Cruz-Campos,  
Juan Carlos de la Cruz-Campos, María Belén Cueto-Martín,  
María García-Jiménez and María Teresa Campos-Blasco  
*University of Granada,  
Department of Physical Education and Sport,  
Spain*

### 1. Introduction

A number of different methodological approaches have been used to describe the inciting event for sports injuries. These include interviews of injured athletes, analysis of video recordings of actual injuries, clinical studies (clinical findings of joint damage are studied to understand the injury mechanism, mainly through plain radiography, magnetic resonance imaging, arthroscopy, and computed tomography scans), in vivo studies (ligament strain or forces are measured to understand ligament loading patterns), cadaver studies and simulation of injury situations, and measurement/estimation from "close to injury" situations. This chapter describes mathematical modeling approach and assesses its strengths and weaknesses in contributing to the understanding and prevention of sports injuries. This chapter demonstrates the relationship between structural measures and lower limb injuries.

Sports injuries can affect any and all parts of the body depending on the particular repetitive movement performed just like any repetitive motion injury. While there are factors that raise the risk of injury, there are also elements that predispose athletes to sports injuries. Rehabilitation and preventative efforts should be centered on a thorough knowledge of risk factor etiology as well as knowledge of how such factors contribute to sports injuries.

In most epidemiological studies directed toward identifying major sports injury causation factors, injured athletes have been compared with uninjured athletes through single variable techniques. However, many of the factors highlighted later in this paper through these analytical techniques either interact or are interrelated.

Multivariable statistical techniques have also been used to detail risk factor interaction (Mechelen, 1992), such as discriminatory analyses and stepwise logistic regression (Dixon, 1993). In this chapter we will identify potential predictive factors that can be used in logistic regression equations, the basic concepts of this mathematical study, and equations that have been developed to what they are today.

### 2. Predictive factors of sports injuries

Predictive factors of sports injuries are biological variables and the relations between them that can be indicators for creating a health profile or diagnosis. For example, weight can be a

predictive factor of diabetes, arteriosclerosis, and other metabolic illnesses. It is even more useful when associated with height, BMI, and waist-hip ratio since it can then be used in predicting hypertension, myocardial infarction, diabetes, and strokes. In order to effectively predict health complications, the WHO recommends using anthropometry to monitor risk factors of chronic diseases and to perform studies that define the association between the aforementioned factors and specific outcomes, such as arterial hypertension. Predicting factors of sports injuries can be grouped into two types of factors: Intrinsic factors and extrinsic factors.

## **2.1 Extrinsic factors**

Sports injuries are most commonly caused by poor training methods; structural abnormalities; weakness in muscles, tendons, ligaments; and unsafe exercising environments. The most common cause of injury is poor training. For example, muscles need 48 hours to recover after a workout. Increasing exercise intensity too quickly and not stopping when pain develops while exercising also causes injury.

The most common cause of sports injuries is improper training whether from a technical or tactical point of view or simply training that is poorly planned and executed (Shaffer, 2006). The athlete exposes him or herself to possible sports injuries without adequate preparation for: exposure to potential danger, the playing position or type of activity, the duration of the competition or league, competition level, time dedicated to training and to rest. These variables can be quantified and turned into predictive factors (Ferrara, 2007). Among such:

**2.1.1** Poor physical condition due to inadequate training (Mechelen, 1996).

**2.1.2** Abrupt increments of training intensity or training load, resulting from overuse and overstress. These injuries tend to appear in underdeveloped locomotion devices, caused by unvaried and unbalanced sports practice especially after training intensification or excessive training. Aerobic training increments between 55% and 75% do not transmit negative effects and do not result in injury risk (Tate, 1995), although disproportionate increases in intensity can provoke anxiety and states of distress.

**2.1.3** Premature competition and quasi-adult training performed by a child or adolescent expose the athlete to injuries from excessive force when pulling or pushing.

**2.1.4** Resistance training, specifically in adolescents, without the appropriate battery of tests to identify the actual state of regulating and homeostatic mechanisms.

**2.1.5** Performing new or unfamiliar exercises. This is common at the beginning of seasons, as well as upon introducing changes in the overall training regime. The same happens when sessions of active rest are planned in which the athlete reflects on unknown sports activities.

**2.1.6** Environmental factors and atmospheric conditions, especially when weather conditions vary unexpectedly or unpredictably, the time of day and the season of the year (Mechelen, 1996).

**2.1.7** The type and integrity of the playing surface, playing surface incline.

**2.1.8** Game mediation: Judges and referees and the official game mediation. Competition among equals is typically promoted in most sports.

**2.1.9** Sports and training equipment. Type and quality of protection, type of footwear, thermal and isothermal clothing. Before using a specific type of shoe, the athlete should know the structural morphology of his or her own feet. Advances in sports injury prevention technology has obligated sports brands to offer a wider variety of models, each model having specific characteristics that complement and correct potential foot dysmorphia.

**2.1.10** Equipment protection, as well as player protection such as shin guards, ankle support, orthotic devices, mouth guards, helmets, prophylactic tape, etc...

**2.1.11** Methodological training development and the level of physicality. Extrinsic factors associated with exposure to injurious situations are: potential dangers, playing positions, competition length, competition level, training time, training frequency, rest intervals or the frequency of exertion, weekly distance run or the number of jumps, hits, throws and impacts, number of trainings per week, training speed, number of competitions per year, absence of regularity in training, etc. All having to do with requirements for the level of physicality needed for different sports activities.

**2.1.12** Prior years of sports training and experience under certain competition and training conditions.

**2.1.13** Inadequate warm-up, whether insufficient or excessive. In some sports, the athlete is required to warm up over such a long period of time that he or she loses concentration on his or her movements and they consequently become ineffective.

**2.1.14** Mastery of a sports technique, technical ability, skill and quality that contributes to the effectiveness of a movement.

**2.1.15** Mental and psychological conditions: Intelligence and creativity, motivation and discipline that influence the precision of technical execution, skill level, previous experiences and the necessity of some athletes to take risks.

**2.1.16** Intrinsic characteristics of common movements in sports activities, linked to structural, biomechanical, and functional characteristics of the athlete. In basketball, the height of the players, the number of jumps, sprints, stops, turning jumps, and backwards jumps can be determining factors in the occurrence of injuries (Grubbs, 1997; Shambaugh, 1991).

**2.1.17** Nutritional and hydroelectric imbalances. (The most common cause of sudden death in marathons is overhydrating).

**2.1.18** The type of sport, notably high-risk and contact sports as well as sports performed when unbalanced or in which unbalancing equipment is used.

**2.1.19** Incorrect playing, inattention to game rules with an excess of rough play and the absence of fair play.

**2.1.20** And, of course, muscular fatigue that stems from technical errors in execution and leads to injury.

## **2.2 Intrinsic factors**

Everyone's bone architecture is a little different, and almost all of us have one or two weak points where the arrangement of bone and muscle leaves us prone to injury.

**2.2.1** Age. There is an increase in the occurrence of injuries in children and adolescents' locomotion devices when they try to perform more ambitiously in hopes of improving their short-term performance. As age and competition level increase, so increases the risk of injury (Inklaar, 1996).

Nonetheless, in many studies, age is not a factor of predisposition, except when it relates to increased speed and distance in training that is significantly greater than that of an older athlete.

In children, the most frequent factors of predisposition to injuries are:

**2.2.1.1** Intrinsic causes in children

**2.2.1.1.1** Muscle tendon imbalance related to strength deficit, excessive flexibility and scant muscle volume.

2.2.1.1.2 Biomechanical alterations, curved, flat feet, femoral anteversion, and the genu valgum that promotes and increased deviation of the Q angle, lumbar hyperlordosis, and length difference between the lower limbs.

2.2.1.1.3 Abrupt weight modifications from growth, since injuries tend to happen from overexertion in growth zones (epiphysitis) when a sudden increase in bone length occurs without parallel adaptation of muscle tendon units.

2.2.1.2 Extrinsic causes in children

2.2.1.2.1 Derived from training errors, high-intensity and long-duration training involving frequent use of still developing structures.

2.2.1.2.2 Planning children's training as though it were adult training, modifying only the workload or volume.

2.2.1.2.3 Technopathies derived from incorrect use of footwear, overly large equipment for the child, hard surfaces, accessories and clothing that are generally inappropriate.

**2.2.2 Sex.** Not a determining risk factor per se, although there are substantial anatomical differences which, in women, are: a wide pelvis, a more pronounced Q angle, greater pelvis anteversion and greater flexibility (Plisky, 2007). Women's levels of training quality and quantity tend to be less intense than that of men due to women's lower muscle mass.

**2.2.3 Structural, neuromuscular conditions that affect athletic performance.** Especially noteworthy conditions are:

2.2.3.1 The alteration of axis lines in the rachis, lumbar curvature, and lower and upper limbs.

2.2.3.2 Lower limb dissymmetry greater than 1 cm.

2.2.3.3 Muscular imbalance: muscular hypotonia and hypertonia.

Genu varum and genu valgum, along with an accentuated Q angle, the genu recurvatum and a smaller intercondylar notch favor the occurrence of injuries in the knees, particularly, cruciate ligaments in women (Shambaugh, 1991). This and the spinal column become real limitations on physical exercise. An excessive Q angle, as well as femoral anteversion favor the occurrence of injuries (Heiderscheit, 1999). The intrinsic causes of jumper's knee, can be sought in the mechanical properties of tendons (resistance, elasticity and extensibility) rather than in morphological or biomechanical abnormalities of the knee extensor mechanism (Ferretti, 1986). Athletes with jumper's knee demonstrated better performance in jump tests than uninjured athletes, particularly in ballistic jumps involving eccentric force generation (Lian, 1996). Bilateral patellar tendinopathy may have a different etiology from unilateral pathology (Gaida, 2004).

2.2.3.4 Functional instability and muscular imbalances in the ankle, ligamentous laxity, peroneal musculature weakness that can diminish control of excessive ankle inversion (Arnold, 2006)). The arch index is a substitute for quantifying foot structure. High-arched runners are at a greater risk of foot injury. Low-arched runners risk soft tissue damage and knee injury (Howard, 2006). While the pronated foot is implicated as a risk factor for **sports injury** in some studies, others suggest that a supinated foot posture increases the risk of overuse lower limb **injuries**. Athletes in a given **sports** discipline may tend to have a similar foot morphology, which varies from that observed elsewhere. Further, the foot morphology that is beneficial for performance in a **sport** may be detrimental with regard to **injury**. (Cain, 2007). Mark (2006) suggest that there are certain factors, including foot pronation, **sport**, and a history of this condition, that are associated with an increased risk of exercise-related leg pain. However, according Barnes (2008) no definitive conclusions can be drawn relating foot

structure or function to an increased risk of tibial stress injuries. Extremes of foot types are likely to pose an increased risk of tibial stress injuries compared to normal arched feet.

2.2.3.5 Joint laxity is still the subject of much debate, although muscular elasticity and flexibility programs are recommended to increase ligaments' defense. (Barber Foss, 2009; Kraemer, 2009)

2.2.4 Warm-up and stretching before competitions (Herbert, 2002, Andersen, 2005).

2.2.5 Height. Taller males are at risk of injury, using single variable logistic regression adjusted to age (Walter, 1989).

2.2.6 High BMI (<19,5 and > 27) relative to the sports activity in question. Simple anthropometric measurements, weight, and age can be effective indicators of future injuries (Backe, 2009). Rose (2008) was found out that students with body mass index (BMI) in the 50th to 90th percentiles had the greatest risk of sport injury. It is concluded that factors like location of residence, ethnicity, and BMI were predictors of sports injuries in adolescents. Men with a waist girth greater than 83 cm seem to be at greater risk of developing patellar tendon pathology. There may be both mechanical and biochemical reasons for this increased risk (Malliaras, 2007).

2.2.7 Deterioration of the senses, such as reduced peripheral vision, myopia or hyperopia, can increase the risk of injury when mistakenly judging the location of teammates, of opponents, the position of the ball or other obstacles.

2.2.8 Somatotype or constitutional type.

2.2.9 Strength and constitutional resistance achieved through training, as well as muscle tone and joint stability. The right- and left-hand power was higher for injured athletes in some sports (Dane, 2002).

2.2.10 Basic conditional qualities such as balance, agility, speed and coordination.

2.2.11 Reaction time and timing.

2.2.12 Physical maturity and posture alignment.

2.2.13 Previous injuries and incomplete recovery from the same before returning to train or compete at the desired tempo and intensity. In these cases, the causes of the injuries may persist, the healed tissue might not work with the required efficacy or the injury might not completely heal. Through a multiple logistic regression analysis, Walter (1989) demonstrated that previous injuries are one of the most indicative factors.

2.2.14 Previous and persisting systemic illness, general and local inflammation, chronic illness, rheumatic diseases, and connective tissue diseases, as well as dental cavities and tonsil stones.

2.2.15 Mental and psychological conditions: intelligence, creativity, motivation, discipline, level of distress, previous experience in the same sports activity, the need to take risks, excessive bravado, fervor, strict adherence to rules and fair play.

### 3. Predictive factors of injuries

When an injury occurs, biomechanical, kinematic, and body composition analyses tend to provide more predictive information than the analyses focused on training intensity, resistance, muscle tone, agility, physical maturity, previous injuries or training methods.

Unevenness in the length of lower limbs, misalignments, anatomical abnormalities, club foot, genu valgum, support type, or posture defects are typically factors cited as injury predictors.

Footprints have also been examined: the average arch, the foot's plantarflexion and dorsiflexion, excessive pronation, as well as the quadriceps' Q angle.

In this chapter we will also delve into constitutional defects in regards to an ideal constitution.

With the exception of the case of athletics, Watson (1987) states that constitutional and postural defects during practice, as opposed to during sports activities, have not received the attention they deserve. Watson points out the clear relation between postural defects and the risk of sports injuries, although it is difficult to prove and establishes a clear relation between foot anomalies and decompensation in the transmission of force in lower limbs, and future repercussions of injury depending on the player's dysmorphia.

### 3.1 The relationship between lower limb structure and sports injuries

Common predisposing factor in injuries to the ankles, legs, knees, and hips include:

Bilateral weight and structural symmetry, Quadriceps and calf girth, patella alta, a kneecap that's higher than usual, Q-angle of the knee (high Q angle: kneecap displaced to one side, as with knock knees), Forefoot varus, Rear foot valgus, true and apparent leg length, uneven leg length, excessive pronation (flat feet), cavus foot (over-high arches), bowlegged or knock-knee alignment.

**Uneven leg length may lead to awkward running and increases the chance of injury**, but many people with equal-length legs suffer the same effects by running on tilted running tracks or along the side of a road that is higher in the centre. The hip of the leg that strikes the higher surface will suffer more strain.

**Pronation is the inward rolling of the foot after the heel strikes the ground**, before the weight is shifted forward to the ball of the foot. By rolling inwards, the foot spreads the shock of impact with the ground. If it rolls too easily, however, it can place uneven stress on muscles and ligaments higher in the leg.

While an overly flexible ankle and foot can cause excessive pronation, a too-rigid ankle will cause the effects of cavus foot. Although the arch of the foot itself may be normal, it appears very high because the foot doesn't flatten inwards when weight is placed on it. Such feet are poor shock absorbers and increase the risk of fractures higher in the legs.

Bowlegs or knock knees add extra stress through knees and ankles over time, and may make ankle sprains more likely.

Other structural conditions that make sports injuries more common include **lumbar lordosis**.

Having some muscles that are very strong and others that are weak can lead to injury. If your *quadriceps* (front thigh muscles) are very strong, it can increase the risk of a stretched or torn hamstring (rear thigh muscle). Tight iliotibial bands may be the cause of knee pain for many athletes in running sports.

**Overuse injuries are caused by repeated, microscopic injuries to a part of the body**. Many long distance runners experience overuse injuries even after years of running. For road runners, the surface is hard and sometimes uneven, and the running movements are repetitive. In addition, there are usually both up- and downhill elements, and these increase the stress on tendons and muscles in the lower leg. You will more likely develop running injuries if you wear the wrong shoes or sneakers. You should use footwear that doesn't allow side-to-side movement of the heel, and that adequately cushions the foot.

Barnes (2008) have not found definitive results that can confirm that constitutional defects are risk factors for injury, while Ferretti (1986) demonstrates that 78% of knee injuries and 50% of spinal column and ankle injuries are closely related to anatomical alteration in static and dynamic foot postures. Regarding these constitutional defects, females could consider themselves at risk due to having a greater articular laxity and less muscle tone, although, at

the same time possessing greater coordination, laterality and body outline. Women also have certain anatomical aspects that can contribute to a greater disposition to injury: a wider pelvis, a greater femoral anteversion, less muscle development of the vastus internus in the quadriceps, a smaller intercondylar notch, a greater tendency towards genu valgum and ligament laxity, external tibial torsion and a higher number of misalignments in lower limbs.

#### 4. Logistic regression equations

The purpose of regression techniques is two-fold:

1. **To estimate the relation between two variables** while taking the presence of other factors into account
2. **To construct a model that allows for the prediction of the value of the dependent variable** (in logistic regression, the probability of success) for specific values of a predicted group of variables

##### 4.1 The concept of logistic regression

The benefit of logistic regression no doubt comes from its capacity to analyze clinical and epidemiological research data. The primary objective that this technique accomplishes is modeling how the presence, or absence, of diverse factors and their values influence the probability of the, typically dichotomic, occurrence of an event. This technique can also be used to estimate the probability of the occurrence of an event with more than two (polytomous) categories.

These sorts of situations are approached using regression techniques. Nonetheless, lineal regression methodology is not applicable since the outcome variable only provides two values (we will focus on the dichotomic case), such as the presence/absence of a knee sprain, or the presence/absence of injury.

If we classify the value of the outcome variable as 0 when the event does not occur (the absence of a knee sprain) and as 1 when it does occur (the athlete sprains his or her knee), and we look to calculate the possible relation between the occurrence of a sprained knee and, for example, the difference in the thickness of both thighs (considered a possible risk factor), we could fall into the temptation of using a linear regression:

$$\text{Knee sprain} = a + b * [\text{difference in thigh thickness}] \quad (1)$$

And, based on our data, gauge the coefficients  $a$  and  $b$  of the equation through the normal procedure of least squares. However, although this is mathematically possible, we arrive at nonsensical results; upon calculating the resulting equation for different values of thigh thickness, we will obtain results that generally differ from 0 and 1, while the only results actually possible in this case are 0 and 1. Since this restriction is not imposed in lineal regression, the outcome can theoretically take on any value.

If we use  $p$  as the dependent variable of probability that an athlete suffers a knee sprain, we can build the following equation:

$$\text{Ln} \frac{p}{1-p} \quad (2)$$

now we do have a variable that can take on any value, and we can therefore propose a traditional regression equation in order to find that value:

$$\ln \frac{p}{1-p} = a + b [\text{difference in thigh thickness}] \quad (3)$$

which, with a slight algebraic manipulation, can be turned into:

$$\text{Injury probability} = \frac{1}{1 + e^{-[-a - b - (\text{difference in thigh thickness})]}} \quad (4)$$

And this is exactly the kind of equation known as a logistic model, where the number of factors can be greater than one. Therefore, in the denominator exponent, we could have:

$$b1.\text{difference in thickness} + b2.\text{age} + b3.\text{sex} + b4.\text{height} \quad (5)$$

#### 4.2 Logistic model coefficients as risk quantifiers

One of the factors that make logistic regression so interesting is the relation that logistic model coefficients preserve with a risk quantification parameter known in the field as an “**odds ratio**” .

The odds associated with an event is the quotient of the probability of occurrence given the probability that it does not occur:

$$\text{Odds Ratio} = \frac{p}{1-p} \quad (6)$$

with  $p$  being the probability of occurrence. Therefore, we can calculate the odds of an injury occurrence when the difference in thigh thickness is equal to or greater than a specific quantity, which determines how much more probable it is that an injury occurs than if it were not to occur in this situation. Likewise, we could calculate the odds of an injury occurrence when the difference in thigh thickness is less than that same figure. If we divide the first odds by the second, we will have calculated an odds quotient, or an odds ratio, which in some way quantifies how probable the occurrence of an injury is when the difference in thickness is greater than a specific figure (first odds) relative to when the difference in thickness is less. The notion being measured is similar to what we find in the **relative risk**, which corresponds to the probability quotient that an injury occurs when a specific factor is present (difference in thickness) compared to when it is not. In fact, when the prevalence of the event occurring is low (<20 %), the odds value ratio and the relative risk are very similar; but such is not the case when the occurrence of the event is quite common, a fact that is often ignored.

$$\text{Relative Risk} = \frac{\text{Probability of Injury the presence of the risk factor}}{\text{Probability of Injury the absence of the risk factor}} \quad (7)$$

$$\begin{aligned} \text{Absolute risk Increase} = & \\ & (\text{post test probability if risk factors is present}) \\ & - (\text{post test probability if risk factors is present}) \end{aligned} \quad (8)$$

If we have a dichotomic factor in the regression equation, for example if the subject is not a jumper, the  $b$  coefficient of the equation for this factor is directly related to the odds ratio **OR** of being a smoker compared to not being one:

$$OR = \exp(b) \quad (9)$$

where  $\exp(b)$  is a measurement that quantifies the risk presented when the corresponding factor is present compared to when it is not, assuming that the rest of the model's variables remain constant.

When the variable is numerical, for example, age or body mass index, it is a measurement that quantifies the change in risk when a variable changes its value while the rest of the variables remain constant. Inasmuch, the odds ratio that, in theory, moves from age  $X_1$  to age  $X_2$ , with  $b$  being the coefficient that corresponds to age in the logistic model is:

$$OR = \exp[b * (X_2 - X_1)] \quad (10)$$

This is a model in which the increase or decrease of risk is proportional to the change in one factor's value to another. In other words, it is proportional to the difference between the two values, but not to the starting point, meaning that the change in risk, in the logistic model, is the same when we move from 20 years old to 30 years old as when we move from 40 to 50.

When the variable's coefficient  $b$  is positive, we obtain an odds ratio greater than 1 that therefore corresponds to a risk factor. On the other hand, if  $b$  is negative the odds ratio will be less than 1 and will correspond to a non-risk factor.

$$\text{Pre-test odds} = \frac{\text{pre-test probability of injury}}{1 - \text{pre-test probability of injury}} \quad (11)$$

$$\text{Pre-test odds} = \text{pre-test odds} \times \text{positive likelihood ratio} - \text{negative-likelihood ratio} \quad (12)$$

Where

$$\text{positive - likelihood ratio} = \frac{\text{sensitivity}}{(1 - \text{specificity})} \quad (14)$$

$$\text{negative - likelihood ratio} = \frac{(1 - \text{specificity})}{\text{sensitivity}} \quad (15)$$

$$\text{Post - test probability} = \frac{\text{post - test odds}}{(\text{post-test odds} + 1)} \quad (16)$$

#### 4.3 Qualitative variables in the logistic model

Given that the employed methodology for calculations with the logistic model is based on using quantitative variables, the same way as in any other regression process, it is incorrect that qualitative variables are used in regression processes, whether nominal or ordinal variables.

Assigning a number to each category does not solve the problem since the physical exercise variable has three possible answers: sedentary, sporadically performing exercise, frequently performing exercise; and we assign the values 0, 1, 2, respectively, to these variables. But then, performing frequent exercise has twice the value of performing exercise sporadically, which makes little sense. Even more absurd would be if a nominal variable, for example

civil status, did not have any ordering relation among the outputs. The solution to this problem is to create as many dichotomic variables as the number of outputs. These new variables, artificially created, are called “dummy”, or indicator, internal, or design variables. Therefore, if the variable in question produces exposure data with the following outputs: *Never ran*, *Ex-runner*, *Runs less than 10 kilometers per day*, *Runs 10 or more kilometers per day*, we have 4 possible answers from which we will construct 3 dichotomic internal variables (values 0,1) with different possibilities for codification that lead to different interpretations. The most frequent of which is the following:

	I1	I2	I3
Never ran	0	0	0
Ex-runner	1	0	0
Runs less than 10 km per day	0	1	0
Runs 10 or more km per day	0	0	1

Table 1. Design variables.

In this type of codification the regression equation's coefficient for each design variable (always transformed with the exponential function), corresponds to the odds ratio for this category given the reference level (the first output). In our example, it quantifies how the risk changes given the situation of never having run. There are other possibilities, among which we will highlight an example with a qualitative variable and three outputs:

	I1	I2
Output 1	0	0
Output 2	1	0
Output 3	1	1

Table 2. Qualitative variable and three outputs.

With this codification, each coefficient is interpreted as an average of the change in risk upon moving from one category to the next. In the event that a category cannot naturally be considered a reference level, for example blood group, a possible classification system is:

	I1	I2
Output 1	-1	-1
Output 2	1	0
Output 3	0	1

Table 3. Classification system of category not natural.

Where each coefficient of the indicator variables has a direct interpretation as a change in risk regarding the average of the three outputs.

**4.4 How to present logistic regression results**

It is common to present logistic regression results in a table wherein each variable will be shown with a coefficient value, its standard error, a parameter (labeled *chi<sup>2</sup> Wald*), which allows us to check if the coefficient is significantly different from 0 and check the *p* value for this context. It also allows us to check the odds ratio of each variable, together with its confidence interval for 95% reliability.

Term	Coeff.	Stand. Err.	chi <sup>2</sup>	p	Interpretation
Indepen.	-1.2168	0.9557	1.621	0.2029	NO
Age	-0.0465	0.0374	1.545	0.2138	NO
Race *			* 5.684	0.0583	Almost (p < 0.1)
Race 1	1.0735	0.5151	4.343	0.0372	p < 0.05
Race 2	0.8154	0.4453	3.353	0.0671	Almost (p < 0.1)
Runner	0.8072	0.4044	3.983	0.0460	p < 0.05
Injury	1.4352	0.6483	4.902	0.0268	p < 0.05
Dissymmetry	0.6576	0.4666	1.986	0.1587	NO
Q Angle	0.8421	0.4055	4.312	0.0379	p < 0.05
Thigh Thickness	1.2817	0.4621	7.692	0.0055	p < 0.01

Table 4. Example of Logistic Regression Presentation.

Variable	Odds ratio	OR < 95%	OR > 95%
Age	0.95	0.89	1.03
Race 1	2.93	1.07	8.03
Race 2	2.26	0.94	5.41
Runner	2.24	1.01	4.95
Injury	4.20	1.18	14.97
Dissymmetry	1.93	0.77	4.82
Q Angle	2.32	1.05	5.14
Thigh Thickness	3.60	1.46	8.91

Table 5. Odds Ratio.

**4.5 Goodness of fit**

As long as we are dealing with a regression model, it is fundamental that the model be checked for an appropriate adjustment to the data used in the calculation before drawing conclusions (Bender, 1996).

In the case of logistic regression, a rather intuitive idea is to calculate the probability of an event, the occurrence of an injury or knee sprain in our case, for all athletes from the sampling. If the goodness of fit is acceptable, one would expect a high probability value to

be associated with the presence of an injury, and vice-versa, if the calculated probability value is low, one would likewise expect the absence of injury.

This intuitive idea is formally realized through the Hosmer-Lemeshow test, that basically consists in dividing the range of probability in deciles of risk (which would be injury probability  $\leq 0.1$ ,  $\leq 0.2$ , and so forth up to  $\leq 1$ ) and calculating the distribution of both injured athletes as well as uninjured athletes that are calculated in the equation and actually observed. These distributions, both calculated and observed, contrast with each other through a  $\chi^2$  test. In the final presentation of logistic regression data, a goodness of fit test should be included as well as a commented conclusion drawn from the same test. With these, the Hosmer-Lemeshow test would be more illustrative than the mere obtained distribution values.

## 5. Logistic regression equation and logistic regression analysis

Despite the fact that accidents are unavoidable in sports, injury prediction and prevention is a practical aspect of sports medicine considered to be the best treatment. Regression models encompass mathematical techniques that deal with measuring the relation between an outcome variable and predictive variables. When the outcome variable is continuous, the preferred model is logistic regression. However, when the outcome variable is dichotomic (injured/not injured) and the object of study is the relation between this and one or more predictive variables (right Q angle, left Q angle, the difference in thigh thickness, lower limb dissymmetry, age, sex, hours of training, kilometers run, etc...) the chosen regression model is a simple logistic regression model (for one factor) or a multiple logistic regression model (for more than one factor).

Therefore, the logistic regression analysis technique is used when it is suspected that one of the values of specific categorical variables depends on a series of predictive or independent variables, along with the goal of finding a mathematical function that expresses such a relation.

When the goal is to calculate the relation or association between two variables, the regression models allow for the consideration that there may be other factors that affect this relation.

So, if the possible relation between lower limb dissymmetry and the probability of suffering a knee injury is being studied as a risk factor, that relation can be different if other variables are taken into account such as age, sex, or body mass index. Because of this, these factors could be included in a logistic regression model as independent variables in addition to dissymmetry. In the resulting equation when considering *DISSYMMETRY*, *AGE*, *SEX*, and *BMI* as independent variables, the *exp* (coefficient of the equation for *DISSYMMETRY*) gives us the **adjusted or controlled odds ratio for the rest of the factors**, given the data for *DISSYMMETRY*.

The other variables, in addition to the interest factor (in this example *AGE*, *SEX*, *BMI*), are called by several names: control variables, external variables, covariants, or **confounding variables**.

### Interaction

When the relation between the factor being studied and the dependent variable is modified by the value of a third variable, we are then dealing with interaction. In our example, we assume that the probability of suffering a **sports injury** increases with age when there is lower limb dissymmetry. In this case we decide that there is an interaction between the variables of *AGE* and *DISSYMMETRY*.

If we focus only on the logistic model exponent, without considering interaction, we would have:

$$-b_0 - b_1 * DISSYMETRY - b_2 * AGE \quad (17)$$

If we want to consider the interaction between INJURY and AGE, the model changes:

$$-b_0 - b_1 * DISSYMETRY - b_2 * AGE - b_3 * DISSYMETRY * AGE \quad (18)$$

If the variable for *DISSYMETRY* is dichotomic (values 0 and 1), the relation between *INJURY* and *DISSYMETRY* will end up quantified by  $b_1$  in the first model while in the second...

$$-(b_1 + b_2 * AGE) * DISSYMETRY \quad (19)$$

In other words, the relation is modified in function of the value of *AGE*.

## 5.1 Precautions

The wide availability of programs that allow access to sophisticated statistical tests can lead to the improper and merely mechanical usage of these tests. Regression models require that the model constructor possess at least a minimal knowledge of the model's underlying philosophy, as well as not only a knowledge of the advantages of this technique, but also of its problems and shortcomings. The use of mathematical processes often convinces us that we are observing "objective" results, and to an extent this is true. However these techniques also carry an intrinsic subjectivity from the selection of a mathematical model to the selection of the variables inserted in that model

### 5.1.1 Independent variable and probability direction

One of the first considerations we must take into account is that the relation between the independent variable and the event probability doesn't change direction. In such a case, the logistic model doesn't work for us. This is something that does not typically occur in clinical studies, but because of that same fact, it is easier to ignore when it does occur.

A very clear example of this situation arises when we evaluate the probability of an athlete's sports injuries in relation to the age when he or she first began sports competitions. Up to a certain age, the probability can increase as the age at which the athlete began competing is earlier. And starting from a mature age, the likelihood of injury also increases compared to the older age at which an athlete competes. In this case, a logistic model would be inadequate.

### 5.1.2 Collinearity

Another problem that may arise in regression models, and not only logistic models, is that the variables involved may be correlated, which would lead us to a nonsensical model and therefore to some values of the coefficients that cannot be interpreted. This situation, with correlated independent variables, is called collinearity.

In order to understand it, let's look at an extreme case in which the same variable is introduced in the model twice. We would then have:

$$\exp (-b_0 - b_1 * X - b_2 * X) \quad (20)$$

or

$$\exp [-b_0 - (b_1 + b_2) * X ] \quad (21)$$

Where the sum of  $b_1+b_2$  allows infinite possibilities when the value of a coefficient is divided into two addends, and therefore the calculation obtained from  $b_1$  and  $b_2$  doesn't make sense.

An example of this situation could be given if we include variables such as the length of the lower limbs and the length of the calves in the equation, two variables that are closely correlated.

### 5.1.3 Sample size

As a basic rule, it is necessary to have at least 10 participants, or  $(k + 1)$  cases to estimate a model with  $k$  independent variables; in other words, at least 10 cases for each dependent variable (the probability of the event).

It is useful to point out that the qualitative variables appear as  $c - 1$  variables in the model, when constructing the corresponding internal variables based on the qualitative variables.

### 5.1.4 Model selection

When talking about models that can be multivariable, an interesting topic is how to choose the best set of independent variables to include in the model (Tsigilis, 2005)

The definition of the "best" model depends on the type and objective of the study. In a case where something will be predicted, the best model would be one that produces the most reliable predictions. And in a case where the relation between two variables is being calculated (correcting the effect of other variables), the best model will be one that obtains the most precise calculation of the coefficient of the variable in question. This is often forgotten and leads to completely different model strategies. Therefore, in the second case a covariant with a statistically significant coefficient, but whose inclusion in the equation does not modify the value of the coefficient of the variable in question, will be excluded from the equation since it doesn't deal with the confounding factor: the relation between the variable in question and the probability are not modified if that variable is taken into account. However, if the outcome of a predictive model is included in the equation, then we look for more reliable predictions.

### 5.1.5 Types of differences

Whenever data is analyzed, it is important to distinguish between numerical differences, statistically significant differences, and clinically relevant differences. These three concepts do not always coincide.

### 5.1.6 Number of variables

The first thing one must consider is the **maximum model**, or the maximum number of independent variables that can be included in the equation, while taking their interactions into account when appropriate.

Although there are different processes for choosing a model, there are only three basic mechanisms for doing so: start with only one independent variable and, one by one, add more according to the pre-established criteria (forward-moving process). Or also, starting with the maximum model, eliminate the variables one by one according to a pre-established criteria (reverse-moving process). The third method, called "*stepwise*", combines the two

previous mechanisms and, in each step, a variable already present in the equation can be eliminated or another can be added.

In the case of logistic regression, the criteria for deciding if we should choose a new model or stay with the currently used one at each step is established by the **models' likelihood ratio logarithm**.

**5.1.7 The likelihood equation**

A model's likelihood equation is a measurement of how compatible the model is with the actual outcome data. If upon adding a new variable to the model, the likelihood does not increase in a statistically significant way, then that variable will not be included in the equation.

To evaluate the statistical significance of a particular variable within the model, we will focus on the Wald chi2 value corresponding to the variable's coefficient and on its level of probability.

**5.1.8 Sports monitoring**

To develop this equation it is necessary to perform a prior monitoring of a representative group of athletes taking into account their age, sex, and sport during a sufficiently long observation period that could be called a season. During this period it is crucial to differentiate the subjects into two groups: injured and non-injured.

Consequently, the relation between the different measured variables and the final outcome of injury or no-injury is established.

In order to determine the predictive variables, we should identify those that show significant differences among the two groups, thus establishing the relation between the injury/no injury dependent variable given the distinct anthropometric and sports variables (activity time, training time, team position, etc...).

**5.1.9 Sensitivity, specificity, positive predictive value and negative predictive value**

It is useful to use control techniques to evaluate the fit of the outcome results. With the mathematical equations defined in the logistic regression analysis. The results should be analyzed in all studied subjects, for the studied group of athletes in question, and for a control group of both sexes and differentiating the success rate by sex.

**5.1.9.1 Sensitivity**

Proportion of injured subjects in relation to how many the equation predicted would be injured.

$$Sensitivity(S_n) = \frac{True\ Positive}{(True\ Positive + False\ Negative)} \tag{22}$$

The following table summarizes these calculations:

	POSITIVE TEST (T+)	NEGATIVE TEST (T-)
INJURY PRESENT (I+)	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
INJURY ABSENT (I-)	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Table 6. Sensitivity.

$$Sn = P [T^+ \text{ if } D^+] \quad (23)$$

$$Sn = \frac{TP}{(TP + FN)} \quad (24)$$

### 5.1.9.2 Specificity

Proportion of uninjured subjects in relation to how many the equation predicted would not be injured.

$$\text{Specificity}(Sp) = \frac{\text{True Negative}}{(\text{True Negative} + \text{False Positive})} \quad (25)$$

You can think of specificity as 1- the false positive rate. Notice what the denominator for specificity is the number of healthy players. Using conditional probabilities, we can also define specificity as:

$$Sp = P[\text{Test is negative if Patient is healthy}] \quad (26)$$

$$Sp = P [T^- \text{ if } I^-] \quad (27)$$

The following table summarizes these calculations:

	POSITIVE TEST (T <sup>+</sup> )	NEGATIVE TEST (T <sup>-</sup> )
INJURY PRESENT (I <sup>+</sup> )	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
INJURY ABSENT (I <sup>-</sup> )	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Table 7. Specificity.

$$Sp = P [T^- \text{ if } D^-] \quad (28)$$

$$Sp = \frac{TN}{(TN + FP)} \quad (29)$$

### 5.1.9.3 False positives

Proportion of uninjured subjects in relation to how many the equation predicted would be injured.

### 5.1.9.4 False negatives

Proportion of injured subjects in relation to how many the equation predicted would not be injured.

In order to know the probability of whether or not a subject injures him or herself in relation to the outcome injury ratio, we must know the positive predictive values (PPV) and the negative predictive values (NPV) that should be defined as the following:

**Positive predictive values:** The probability of an athlete injuring him or herself when predicted by the equation. To calculate this we use the equation:

$$PPV = \frac{(S * PL)}{(S * PL) + (FL * PNL)} \tag{30}$$

Where S: Sensitivity. PL: Probability of injury. FP: False Positives. PNL: Probability of non-injury.

The following table summarizes these calculations

	POSITIVE TEST (T <sup>+</sup> )	NEGATIVE TEST (T <sup>-</sup> )
INJURY PRESENT (I <sup>+</sup> )	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
INJURY ABSENT (I <sup>-</sup> )	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Table 8. False negatives.

$$PPV = P [I^+ \text{ if } T^+] \tag{31}$$

$$PPV = \frac{TP}{(TP + FP)} \tag{32}$$

**Negative predictive values:** The probability that the athlete does not injure him or herself when the model has predicted a situation of non-injury. To calculate this we use the equation:

$$NPV = \frac{(E * PNL)}{(E * PNL) + (FN * PL)} \tag{33}$$

	POSITIVE TEST (T <sup>+</sup> )	NEGATIVE TEST (T <sup>-</sup> )
INJURY PRESENT (I <sup>+</sup> )	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
INJURY ABSENT (I <sup>-</sup> )	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Table 9. Negative predictive values.

$$NPV = P [I^- \text{ if } T^-] \tag{34}$$

$$NPV = \frac{TN}{(TN + FN)} \tag{35}$$

Where S: Sensitivity. PL: Probability of injury. FP: False Positives. PNL: Probability of non-injury.

It is always necessary to find false negatives and positives beforehand, as well as the probability of injury or non-injury for each athlete before determining the positive and negative predictive values.

In order to perform this type of calculation, the probability that an individual exhibits the characteristic in question (suffering an injury) is expressed in function of the predictive variable or variables; if we make P the probability, the model is expressed as follows:

$$P = \beta_0 + \beta_1 X \tag{36}$$

Where  $\beta_0$  y  $\beta_1$  are the model parameters and  $X$  is the predictive variable. The probability ( $P$ ) is equal to a constant  $\beta_0$  plus the product of the other constant  $\beta_1$  multiplied by the value of the predictive variable  $X$ .

The coefficient  $\beta_0$  is an independent or constant term and it is the value of the outcome variable's average. The coefficient  $\beta_1$  is the regression coefficient and it is interpreted as the change in the outcome variable's average by the unit of increase of the predictive variable. The change will be an increase if the regression coefficient value is positive and it will be a decrease if the value is negative.

It is possible that once the model parameters are calculated, the substitution of some values of the predictive variable gives way to values that aren't allowed for a probability. This is why one should perform a probability transformation for the probability of showing the characteristics in question. This logit transformation that consists in the logarithmic odd

$\frac{p}{1-p}$  that a characteristic will present itself, is modeled by the following formula:

$$\text{Log} \left[ \frac{p}{1-p} \right] = \beta_0 + \beta_1 X \quad (37)$$

$$\text{The Log} \left[ \frac{p}{1-p} \right] \text{ is called logit (P)} \quad (38)$$

In the logistic regression model, the coefficient  $\beta_1$  is the logarithm of the odds ratio (OR) between two individuals that are differentiated in a unit in terms of the predictive variable. Likewise, by raising  $e$  to  $\beta_1$ , we obtain the OR value between those two individuals.

$$\text{Log (O.R.)} = \beta_1 \quad (39)$$

Or:

$$\text{O.R.} = e^{\beta_1} \quad (40)$$

where  $e$  is the number that serves as the base of the Napierian logarithm, approximately 2.72.

In the logistic regression model,  $\beta_1$  is the OR logarithm between two individuals that are differentiated in a unit in terms of the predictive variable, or likewise, by raising  $e$  to  $\beta_1$ , one obtains the OR value between these two individuals. In the case where  $\beta_1=0$ , it is implied that the  $\text{logit}(P) = \beta_0 + (0) X = 0$ , in other words, it does not change with  $X$ . Or equally,  $\text{O.R.} = e^0 = 1$ , which indicates that the two variables are independent and there is no relation between them. The calculation of  $\beta_1$  is called the logistic regression coefficient.

If we have several predictive variables and we try to study the relation between the outcome variable and the whole set of predictive variables simultaneously, a multiple logistic regression model will be used.

$$\text{Log} \left[ \frac{p}{1-p} \right] = \beta_0 + \beta_1 X + \dots + \beta_p X_p \quad (41)$$

where  $P$  is also the probability of presenting the characteristic in question.

An alternative form of presenting the same model is:

$$P = \frac{e^{(\beta_0 + \beta_1 + \dots + \beta_p X_p)}}{1 + e^{(\beta_0 + \beta_1 + \dots + \beta_p X_p)}} \quad (42)$$

Which allows the calculation of the probability that an individual with certain predictive values will exhibit the characteristic in question.

## 5.2 Logistic regression equations applied to sports

### 5.2.1 Shambaugh injury score

Shambaugh (1991) proposed the first logistic regression equation that would predict injuries occurring within a season with 91% accuracy. The variables initially proposed were the diameter of the thigh, diameter of the calf, the Q angle, ankle dorsiflexion, genu varum and valgum, the difference in supported weight, and the length of the legs. Ankle dorsiflexion and varum were more elevated in uninjured players, and therefore we reject them in the final equation.

When the outcome of the equation was positive, the subject was predicted to be at risk for injury. But one equation with so many relative coefficients and too many collateral effects was, aside from being difficult to design, too complicated to be reliable.

Shambaugh determined that there should only be three dependent variables selected, and since one of the fundamental goals was to find structural asymmetries or imbalances, he opted for the Q angle of the knees and the difference in supported weight between both legs, resulting in the following equation, applicable only to males:

$$\text{SHAMBAUGH injury score (1991)} = \left( \begin{array}{l} \text{imbalance in bilateral weight* .36+right abnormal Q angle*0.48+} \\ \text{left abnormal Q angle*0.86} \end{array} \right) - 7.04 \quad (43)$$

considering the weight imbalance between the right and left leg to be an absolute value and the abnormality of the Q angle for males beginning at 10°.

The value of the Shambaugh score would be directly proportional to the possibility of injury; the higher the score was, the greater the probability of including the athlete in the injury category. There was a 95% success rate. The player that obtained a higher score was also the one who incurred the most serious injury.

Grubbs (1997) studied the Shambaugh score in relation to its predictive value, calculating the statistical values of sensitivity, specificity and positive and negative prediction in men as well as women. Its results were less outstanding, upon inclusion of results for women, the abnormality value of the Q angle in women was set at 15°.

In 2000 he proposed a modification to his injury score, introducing 4 variables instead of the three original ones in his logistic regression equation. The new variable was the squared value of the difference between thigh thickness:

$$\begin{aligned} \text{SHAMBAUGH Injury Score (2000)} = & \\ & \text{Weight Imbalance} * 0.27 + 1.46 * (\text{Difference in thigh thickness})^2 \\ & + 0.22 * \text{Difference in the Q angle of both knees} \\ & + 0.94 * \text{Right abnormal Q angle} - 6.46. \end{aligned} \quad (44)$$

### 5.2.2 Salazar injury score (Salazar, 2000)

Salazar (2000) expanded on the Shambaugh injury score by including data from exposure to injury at practice, training time, and game play time. He used the control sheets created by DeLee (1992):

$$\text{SALAZAR INJURY PROBABILITY SCORE} = \frac{1}{1 + e^{(0.1621 - 0.06344 * \text{average Shambaugh score})}} \quad (45)$$

### 5.2.3 Fernández-de la cruz injury score (Fernández-Martínez, 2008)

$$\text{FERNANDEZ - DE LA CRUZ DE INJURY PROBABILITY SCORE} = \frac{1}{1 + e^{-\left(0.757 * \text{AQI} - 0.647 * \text{DGM}^2\right)}} \quad (46)$$

where AQI is the left knee Q angle and DGM<sup>2</sup> is the squared value of the difference in thigh thickness, demonstrating a 72.9% success rate for injury prediction (positive prediction at 75.68%; negative prediction at 70.73%). This equation is applicable to men as well as to women. The overall percentage of correct classification was 68.6%. The cutoff point (0.5) indicates that the subjects with values equal to or greater than 0.5 would be placed in the “at risk” category, while a value less than the cutoff point would place them in the “reduced risk for injury” category.

## 6. Conclusion

Logistic regression equations allow injury prediction for athletes, risk calculation, and the opportunity for establishing the most effective and appropriate measures to be taken. Its versatility and capacity for being applied to specific sports groups allows personalized attention for each group.

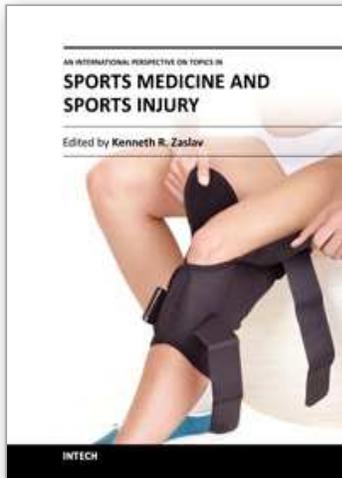
This chapter shows that the logistic regression analysis can be used as a valid method in determining anthropometric parameters related to sports injuries, while providing a reliable and simple method that can be used in the common practice of sports medicine.

## 7. References

- Andersen, J. C. (2005). Stretching Before and After Exercise: Effect on Muscle Soreness and Injury Risk. *Journal of Athletic Training*, Vol. 40 Issue 3, 218-220. ISSN 1062-6050.
- Arnold, Brent L.; Doherty, Carrie L. (2006). Low-Load Eversion Force Sense, Self-Reported Ankle Instability, and Frequency of Giving Way. *Journal of Athletic Training*, Vol. 41 Issue 3, 233- 239. ISSN 1062-6050.
- Backe, S.; Ericson, L.; Janson, S.; Timpka, T. (2009). Rock climbing injury rates and associated risk factors in a general climbing population.. *Scandinavian Journal of Medicine & Science in Sports*, Vol. 19 Issue 6, 850- 856. ISSN 0905-7188
- Barber Foss, Kim D.; Ford, Kevin A.; Myer, Gregory D.; Hewett, Timothy E. (2009). Generalized Joint Laxity Associated With Increased Medial Foot Loading in Female Athletes. *Journal of Athletic Training*, Vol. 44 Issue 4, 356- 362. . ISSN 1062-6050.

- Barnes, A.; Wheat, J.; Milner, C. (2008) Association between foot type and tibial stress injuries: a systematic review. *British Journal of Sports Medicine*. Vol. 42 Issue 2, 93- 98. ISSN 0306-3674.
- Bender, R.; Grouves, U. (1996) Logistic regression models used in medical research are poorly presented. *British Medical Journal*. 313, 628 ISSN: 0959- 8138.
- Bowers, Andrea L.; Spindler, Kurt P.; McCarty, Eric C.; Arrigain, Susana (2005). Height, Weight, and BMI Predict Intra-articular Injuries Observed During ACL Reconstruction: Evaluation of 456 Cases From a Prospective ACL Database. *Clinical Journal of Sport Medicine*: Vol. 15 Issue 1. p. 9-13. ISSN 1536-3724.
- Cain, L.; Nicholson, L.; Adams, R.; Burns, J. (2007) Foot morphology and foot/ankle injury in indoor football. *Journal of Science & Medicine in Sport* . Vol. 10 Issue 5. p. 311-319. ISSN 1440-2440.
- Dane, S.; Can, S.; Karsan, O. (2002). Relations of body mass index, body fat, and power of various muscles to sport injuries. *Perceptual & Motor Skills*: Vol. 95 Issue 1. p. 329-334. ISSN 0031-5125.
- Dixon, W.J. Brown, M.B.; Engelman, L. and Jennerich, R.I. (1993). *BMDP Statistical Software Manual, 1992 Edition, 3 Volume set*. pp 1105- 45. University of California Press. ISBN 978-05-2008-141-3. Berkeley.
- Fernández, A.; De la Cruz-Márquez, J.C.; Cueto-Martín, B.; Salazar Alonso, S.; De la Cruz-Campos, J.C. (2008). Predicción de lesiones deportivas mediante ecuaciones de regresión logística. *Apunts Medicina de l'esport*, 157- 41-44. ISSN: 1886- 6581.
- Ferrara, C.M.; Hollingsworth, E. (2007) Physical Characteristics and Incidence of Injuries in Adult Figure Skaters. *International Journal of Sports Physiology & Performance*, Vol. 2 Issue 3, 282- 291. ISSN 1555-0265.
- Ferretti, A. (1986) *Epidemiology of jumper's knee*. *Sports Medicine*.; 3:289-95. ISSN: 1179-2035.
- Finkelstein, E.A.; Hong Chen; Prabhu, M; Trogdon, J.G.; Corso, P.S. (2007) The Relationship Between Obesity and Injuries among U.S. Adults. *American Journal of Health Promotion*, Vol. 21 Issue 5, 460- 468. ISSN 0890-1171.
- Gaida J, Cook J, Bass S, Austen S.; Kiss, Z.S. (2004) Are unilateral and bilateral patellar tendinopathy distinguished by differences in anthropometry, body composition, or muscle strength in elite female basketball players? *British Journal of Sports Medicine*. 38:581-5. ISSN 0306-3674.
- Gissane, Conor; White, John; Kerr, Kathleen; Jennings, Deanna (2001) An operational model to investigate contact sports injuries. *Medicine & Science in Sports & Exercise*: Vol 33 - Issue 12 - pp 1999-2003. ISSN 0195-9131.
- Grubbs, N.; Nelson, R.T.; Bandy, W.D. (1997) Predictive validity of an injury score among high school basketball players. *Medicine & Science in Sports & Exercise* Oct 1997: Vol. 29 Issue 10. 1279-1285. ISSN 0195-9131.
- Heiderscheit, B.C.; Hamil, J.; Van Emmerik, R (1999) Q-angle influences on the variability of lower extremity coordination during running. *Medicine Science and Sport Exercise*, 31, 1313- 1319 ISSN: 1530- 0315.
- Herbert RD, Gabriel M. (2002) Effects of stretching before and after exercise on muscle soreness and risk of injury: systematic review. *British Medical Journal*. 325:468. ISSN 0959-8138.
- Howard, J.S.; Briggs, D. (2006) The Arch-Height-Index Measurement System: A New Method of Foot Classification. *Athletic Therapy Today*, Vol. 11 Issue 5, 56-57. ISSN 1078-7895.
- Kraemer, R.; Knobloch, K. (2009) A Soccer-Specific Balance Training Program for Hamstring Muscle and Patellar and Achilles Tendon Injuries: An Intervention Study in Premier League Female Soccer. *American Journal of Sports Medicine*, Vol. 37 Issue 7, 1384- 1394. ISSN 0363-5465.

- Inklaar, E.; Bol, E.; Schmikli, S.L.; Mosterd, W.L. (1996) Injuries in male soccer players: team risk analysis. *International Journal of Sports Medicine*, 17, 229- 234. ISSN: 0172- 4622.
- Lian O, Engebretsen L, Ovrebo RV, Bahr, R. (1996) Characteristics of the leg extensors in male volleyball players with jumper's knee. *American Journal of Sports Medicine*; 24:380-5. ISSN 0363-5465.
- Lian O, Refsnes P, Engebretsen L, Bahr, R. (2003) Performance characteristics of volleyball players with patellar tendonopathy. *American Journal of Sports Medicine*, 31:408-13. ISSN 0363-5465.
- Lowry, R.; Lee, S. M.; Galuska, D. A.; Fulton, J. E.; Barrios, L. C.; Kann, L. (2007) Physical Activity-Related Injury and Body Mass Index Among US High School Students. *Journal of Physical Activity & Health*. Vol. 4 Issue 3, 325- 342. ISSN 1543-3080.
- Malliaras, P.; Cook, J.L.; Kent, P.M. (2007) Anthropometric risk factors for patellar tendon injury among volleyball players. *British Journal of Sports Medicine*. Vol. 41 Issue 4. 259-263. ISSN 0306-3674.
- Mark F. (2006) Exercise-Related Leg Pain in Female Collegiate Athletes The Influence of Intrinsic and Extrinsic Factors. *American Journal of Sports Medicine*: Vol. 34 Issue 9. pp. 1500-1507. ISSN 0363-5465.
- Mechelen, V. (1999). Lesiones en el atletismo. In: *Prácticas clínicas sobre asistencia y prevención de lesiones deportivas*. Renström, P.A. Paidotribo Publisher. pp: 479- 506. ISBN: 978-84- 8019- 420-4. Barcelona.
- Meyers, M.C. and Barnhill, B.S. (2004). Incidence, Causes, and Severity of High School Football Injuries on FieldTurf Versus Natural Grass. A 5-Year Prospective Study. *American Journal of Sports Medicine*., 32, 1626-1638. ISSN 0363-5465.
- Plisky, M. S.; Rauh, M. J.; Heidrscheit, B.; Underwood, F.B.; Tank, R.T. (2007) Medial Tibial Stress Syndrome in High School Cross-Country Runners: Incidence and Risk Factors.. *Journal of Orthopaedic & Sports Physical Therapy*, Vol. 37 Issue 2, 40- 47. ISSN: 0190-6011.
- Rose, M. S.; Emery, C.A.; Meeuwisse, W.H (2008) Sociodemographic Predictors of Sport Injury in Adolescents. *Medicine & Science in Sports & Exercise*. Vol. 40 Issue 3, 444-450. ISSN: 0195-9131
- Salazar, S. (2000). Aplicación del Índice de Shambaugh en jugadores/as de baloncesto cadetes y júnior en relación con la exposición práctica. Doctoral Thesis. University of Granada. Spain.
- Shaffer, S.W.; Uhl, T. L. (2006) Preventing and Treating Lower Extremity Stress Reactions and Fractures in Adults. *Journal of Athletic Training*. Vol. 41 Issue 4, 466- 470. ISSN: 1062-6050.
- Shambaugh, J.; Klein, A.; Herbert, J. (1991). Structural measures as predictors of injury in basketball players. *Medicine & Science in Sports & Exercise* , 23, 522- 527. ISSN 0195-9131.
- Tate, A.; Petruzello, S (1995). Varying the intensity of acute exercise: implications for change in affect. *Journal of Sports Medicine and Physical Fitness*, 35, 295- 302. ISSN 0022- 4707.
- Tsigilis, N.; Hatzimanouil, D. (2005) Injuries in handball: Examination of the risk factors. *European Journal of Sport Science*. Vol. 5 Issue 3, 137- 142. ISSN: 1536-7290.
- Walter, S.D.; Hart, L.; McIntosh, J. (1989) The Ontario cohort study of running related injuries. *Archives of Internal Medicine*, 149, 2561- 64. ISSN 0003- 9926.
- Watson, M. Paul, P. (1987) Incidence of injuries in high school track and field athletes and its relation to performance ability. *American Journal of Sports Medicine*, 15, 251-54. ISSN 0363-5465.



## **An International Perspective on Topics in Sports Medicine and Sports Injury**

Edited by Dr. Kenneth R. Zaslav

ISBN 978-953-51-0005-8

Hard cover, 534 pages

**Publisher** InTech

**Published online** 17, February, 2012

**Published in print edition** February, 2012

For the past two decades, Sports Medicine has been a burgeoning science in the USA and Western Europe. Great strides have been made in understanding the basic physiology of exercise, energy consumption and the mechanisms of sports injury. Additionally, through advances in minimally invasive surgical treatment and physical rehabilitation, athletes have been returning to sports quicker and at higher levels after injury. This book contains new information from basic scientists on the physiology of exercise and sports performance, updates on medical diseases treated in athletes and excellent summaries of treatment options for common sports-related injuries to the skeletal system.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Juan Carlos de la Cruz-Márquez, Adrián de la Cruz-Campos, Juan Carlos de la Cruz-Campos, María Belén Cueto-Martín, María García-Jiménez and María Teresa Campos-Blasco (2012). Prediction of Sports Injuries by Mathematical Models, An International Perspective on Topics in Sports Medicine and Sports Injury, Dr. Kenneth R. Zaslav (Ed.), ISBN: 978-953-51-0005-8, InTech, Available from:  
<http://www.intechopen.com/books/an-international-perspective-on-topics-in-sports-medicine-and-sports-injury/prediction-of-sports-injuries-by-mathematical-models>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen