Universidad Miguel Hernández de Elche

DEPARTAMENTO DE PSICOLOGÍA DE LA SALUD

Programa de Doctorado en Deporte y Salud

EPIDEMIOLOGY AND PREDICTIVE MODELS OF INJURIES IN PROFESSIONAL SOCCER

Doctoral thesis

A dissertation presented by
Alejandro López Valenciano

Graduate in Physical Activity and Sports Science

Elche, 2017

El Dr. D. Eduardo Cervelló Gimeno, coordinador del programa de doctorado en Deporte y Salud de la Universidad Miguel Hernández de Elche.

AUTORIZA:

Que el trabajo de investigación titulado: “EPIDEMIOLOGY AND PREDICTIVE MODELS OF INJURIES IN PROFESSIONAL SOCCER” realizado por D. Alejandro López Valenciano bajo la dirección del Dr. D. Francisco José Vera García y del Dr. D. Mark de Ste Croix sea depositado en el departamento y posteriormente defendido como Tesis Doctoral en esta Universidad ante el tribunal correspondiente.

Lo que firmo para los efectos oportunos en

Elche, Octubre de 2017

Fdo.: Eduardo Cervelló Gimeno

Coordinador del programa de doctorado

Universidad Miguel Hernández de Elche

UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

Departamento: Psicología de la Salud

Programa de Doctorado: Deporte y Salud

Título de la Tesis

EPIDEMIOLOGÍA Y MODELOS DE PREDICCIÓN DE LESIONES EN EL FÚTBOL PROFESIONAL

Tesis Doctoral presentada por:

D. Alejandro López Valenciano

Dirigida por el Dr. D. Francisco José Vera García y

Co-dirigida por el Dr. D. Mark de Ste Croix

Los directores **El Doctorando**

Elche, Octubre de 2017

DEPARTAMENTO DE PSICOLOGÍA DE LA SALUD

Programa de Doctorado en Deporte y Salud

EPIDEMIOLOGY AND PREDICTIVE MODELS OF INJURIES IN PROFESSIONAL SOCCER

A dissertation presented by
Alejandro López Valenciano

Directores:
Dr. D. Francisco José Vera García
Dr. D. Mark de Ste Croix

AGRADECIMIENTOS

Desde muy pequeño siempre me han dicho que es de bien nacidos ser agradecidos, y creo que la gente que lea estos agradecimientos me conoce lo bastante bien como para saber que el señor no me ha dado el don de la extraversión. Por ello, no hay mejor manera de reflejar todo lo que pienso que recordando y agradeciendo a todos aquellos que han puesto una miguita en el camino para que pueda presentar esta tesis.

Como no, todo parte de la familia, en el momento en que mis padres pusieron toda su ilusión por tener otro hijo y darle un hermanito a Javi. Mis padres, Isidoro y Cande, son el claro ejemplo de lo que uno quiere ser en la vida, humildes, luchadores y con unos principios que me supieron transmitir desde muy pequeño, humildad, perseverancia y pasión por lo que te gusta. Y además de por las razones económicas obvias para que yo pudiese estudiar, les debo el poder dedicarme a mi PASIÓN, puesto que desde muy pequeño han aguantado y “apoyado” todo lo que hacía. Y no solo eso, ellos pusieron las primeras migas con su esfuerzo en formarme día a día, mi madre siempre atenta y paciente ayudándome con los deberes y mi padre, en lugar de descansar después de todo el día trabajando, lo primero que hacía al llegar a casa era ayudarme a repasar los exámenes. Os quiero!

La otra gran parte de mi familia es lo que se debería definir en los diccionarios como un hermano mayor, mi hermano Javi. Desde muy pequeño ha cuidado de mi mientras mis padres trabajaban, donde él iba allí estaba yo, cosa que hacía, cosa que imitaba yo, y obviamente, suponía más que un incordio y resultado de ello más de una colleja ha volado. Seguramente sea esta la razón de que él sea el más inteligente de los dos. Aun cuando somos polos opuestos en muchos sentidos, siempre ha sido el espejo en el que mirarme, por su empeño y dedicación en lo que se propone, con independencia del esfuerzo y los pesares del camino. Y aunque esos primeros años fueron difíciles, ahora soy yo el que “cuida” de él en el aspecto deportivo, lo que nos ha acercado mucho más, convirtiéndose en la persona de más confianza para mí, y claro está que te ayude a pagar un coche hace mucho.

Mención especial merecen los que no pueden estar aquí presentes, especialmente la persona que marcó un antes y un después en el camino durante mi primer año de carrera, mi tío Fernando. Era todo un referente para mí como profesor de Educación Física y como persona, siempre animándome a buscar mi sueño. Esto me enseñó a valorar lo que es realmente importante en la vida y disfrutar y luchar por lo que verdaderamente quiero. Igualmente, aunque no pude disfrutar lo que hubiese querido de ellos, añoro los primeros años de vida con mis abuelos: Javier, María, Antonio y Ventura. Pequeñas cosas como “dame 20 duros”, “eres más listo que un reloj”, jugar a las cartas o quitarle la gorra a mi abuelo Javier, o la dulzura de mi abuela Ventura son cosas que nunca olvidaré.

¿Pero de dónde surge mi pasión por la actividad física y el deporte? Sinceramente no recuerdo el día ni el motivo porque no vengo de una familia que le guste especialmente el deporte, pero al igual que muchas personas no tienen muy clara su vocación, yo recuerdo monopolizar desde muy pequeño el mando de la tele con el deporte y llegar a casa con menos de 10 años y decirle a mis padres “quiero ser como Don Andrés”. Don Andrés fue mi primer profesor de Educación física y seguramente una de las personas por las que hoy esté aquí. Tras unos años en los que no quedaba muy claro muy futuro y tras la fortuna de encontrarme con una profesora, Maria Angeles Trueba, que supo reconducir mi camino, pude realizar las pruebas de acceso a Ciencias de la Actividad Física y del deporte en Toledo, aunque mi tan conocida potencia de salto negativa no dio para mucho en las pruebas y tuve que reconducir mi camino hacia Elche, hacia la UMH, y que suerte la mía!

Aquí es cuando comienza otra etapa en mi vida, cuando empiezo a ser consciente de lo que realmente son las ciencias del deporte y donde pude disfrutar de años fantásticos con mi grupo “Brasil”, Jacobo, Fliso, Ratika, Flavio, Jony, Leguey, Muñoz… sin los cuales hubiese sido imposible disfrutar tanto. Durante los seis años de carrera y máster fui forjando mi idea de poder formarme y trabajar con profesionales como los que fueron mis profesores, de los cuales estaré siempre agradecido: Pepo, Eduardo, Raúl Reina, Fito, Rafa, Manolo, Francis, Peláez, Celes, Carlos, Diego Pastor, Vicente, David González…

En esos momentos, me decanté por la persona que más admiración me había generado y me sigue generando como persona y como profesional, mi director de tesis, Fran Vera. Recuerdo aquella reunión en quinto de carrera en la que le pedía la oportunidad de poder ayudar en el Centro de Investigación del Deporte (CID) de la UMH y seguir vinculado con el ámbito científico. Él como siempre fue más que sincero y honesto desde el principio, aconsejándome y dejando claras las posibilidades de poder hacerlo. Desde ese día he disfrutado de su aprendizaje, perfeccionista y meticuloso, que me ha hecho ver por qué es un referente a nivel mundial en nuestro ámbito. Además, siempre es el primero en preocuparse de ti, en ayudarte, aconsejarte, y permitirte no cerrar ninguna oportunidad profesional, aun cuando algún disgustillo le he dado con mis “escapadas” al Valencia Open o a la Ruta Quetzal, pero que cobra con sus “decentes” imitaciones sobre mí. Estas “escapadas” también me permitieron conocer gente maravillosa que forma parte de mi vida y que me han marcado de algún modo u otro, María Jesús, Natalia, monis, mi grupo 9 y Zulema, que siempre tiene palabras de ánimo cuando más las necesito.

Si Fran Vera es un ejemplo, no lo menos es el genio de mi otro director de tesis, Fran Ayala. Fran fue la persona que me dio el empujón que necesitaba para convencerme de que lo que estaba haciendo era lo que quería en la vida. Fran, desde su posición de director de tesis, siempre ha sabido ponerse a mi altura y darme total confianza, cercanía y seguridad para que luchemos en ser referentes en nuestro ámbito y no esperar a que las cosas lleguen. Fran ama lo que hace y lo basa todo en ello, como refleja su curriculum y su pundonor diario, haciéndonos mejorar a todos o al menos presionando con frases como “descansa el finde y ya lo haces la semana que viene”, y luego ver un correo diciéndote que el lunes debe estar todo hecho. Pero esto es lo que ha hecho que podamos presentar una tesis de tal calidad, la cual no sería posible sin él. En este apartado también tengo mucho que agradecer a mí co-director, Mark de Ste Croix, el cual me permitió dar un salto de calidad en mi formación académica, con su profesionalidad, cercanía y amabilidad durante mi estancia en el extranjero.

Pero esto no acaba aquí, el camino te lleva a conocer a grandes personas que aman lo mismo que tú, aquellos que viven el deporte como yo en el CID. Cuando entras al CID, lo primero que aprendes es que hay un padre y una madre cariñosamente hablando, que nos ponen en su sitio a todos los “pipiolos”. Pero no solo eso, son los que te ayudan desinteresadamente, los primeros en recibirte con una sonrisa cada mañana y en felicitarte y apoyarte cuando lo necesitas. Gracias Juan Pedro y Dori!. Lo segundo es que te vas a ir de ahí con una familia mucho más grande. Dentro de esa familia está mi grupo, mis referentes y aquellos que no dudan en primar el grupo por el interés propio, BIOMEC. Ahí está controlando y solucionando todos los aspectos tecnológicos el genio de la lámpara, Pepo. Pepo siempre soluciona en un momento cualquier problema tecnológico que parece un mundo, o se saca de la manga un software que te ahorra meses de trabajo. Luego están David, con su ayuda siempre a “tiempo” ;) y el cual me ha enseñado que nunca debes conformarte con lo justo, puesto que siempre habrá algún análisis o proceso para mejorar tu trabajo, y Casto, que quién iba a decirnos donde acabaríamos cuando empezamos juntos ya hace unos años por otro sendero, y del que me alegro que se haya convertido en todo un “jefe” del core. Y como no, mi Yang en el grupo, Iñaki, que no apuntaba muchas maneras cuando llegó el primer día al laboratorio con sus rastas y no quería usar la libreta para aprender a manejar el VICON. Sin embargo, y como con todos los asuntos que trates con él, sabe darle la vuelta y llegar a ganarte, pero no de forma desmerecida, sino todo lo contrario, puesto que da lo que haga falta por sus compañeros y en mi caso ha sido un apoyo fundamental para conseguir esta tesis, sobre todo cuando hace de mi secretaria. No me quiero olvidar tampoco de los que vienen por detrás y que también han puesto su granito de arena en estos últimos meses y los cuales estoy seguro se convertirán en referentes de nuestro campo, Pedro, Belén y Diego.

Dedicatoria personal merecen Miguel y MariPili, puesto que fueron dos pilares importantes en los primeros años en Elche y especialmente en los momentos duros del camino, con sus risas, con sus pequeñas broncas (MP) y sobre todo con palabras de ánimo y apoyo. Ya han pasado muchos años desde que nos conocimos pero conforme pasan los años tengo claro que serán parte de mí toda la vida.

Y como iba diciendo, el CID es una familia que va creciendo poco a poco, y a la que se unen personas que priman la amistad por delante de todo y sin ellos, sería prácticamente imposible escribir estas líneas. Ya fuese de tapas, de fiesta, en cualquier pachanga o escapadita laboral me han permitido desconectar y hacer este tramo mucho más ameno. Gracias Tomás, Héctor, Raúl López, Jaime, Alba, Carla, Choche, Marta, María, Carlos, Juanmi, Alberto, Adri, Diego Pastor, Sarabia, Santi, Ramón, Alicia, Álvaro, Artur, Félix, Java,…Esperemos que nunca se pierda ese buen rollo y podamos hacer muchos cosas juntos.

También agradecer a mi Peña el apoyo importante a la hora de desconectar en el pueblo, y aunque sean pocas las veces que nos vemos, las risas que nos echamos dan para mucho.

Para terminar, agradecer la financiación recibida por parte del Ministerio de Educación, Cultura y Deporte a través del programa de formación del profesorado Universitario (FPU).

Sin más, y como lo mío es un deporte de resistencia, espero que esto solo sea el primer paso de una gran carrera de fondo en la que pueda apoyarme de grandes personas como las que han hecho posible este trabajo.

TABLE OF CONTENTS

LIST OF FIGURES

Figure 5.11. Muscle injuries predictive model, classifier 10.

Figure 5.12. Graphical representation of the first classifier. Prediction nodes are represented by ellipses and splitter nodes by rectangles. Each splitter node is associated with a real valued number indicating the rule condition, meaning: If the feature represented by the node satisfies the condition value the prediction path will go through the left child node, otherwise the path will go through the right child node. The numbers before the feature names in the prediction nodes indicate the order in which the different base rules were discovered. This ordering can to some extent indicate the relative importance of the base rules.

Figure 6.1. Testing procedure.

Figure 6.2. Hamstring strain injury predictive model, classifier 1.

Figure 6.3. Hamstring strain injury predictive model, classifier 2.

Figure 6.4. Hamstring strain injury predictive model, classifier 3.

Figure 6.5. Hamstring strain injury predictive model, classifier 4

Figure 6.6. Hamstring strain injury predictive model, classifier 5.

Figure 6.7. Hamstring strain injury predictive model, classifier 6.

Figure 6.8. Hamstring strain injury predictive model, classifier 7.

Figure 6.9. Hamstring strain injury predictive model, classifier 8.

Figure 6.10. Hamstring strain injury predictive model, classifier 9.

Figure 6.11. Hamstring strain injury predictive model, classifier 10.

Figure 6.12. Example of alternating decision tree.

Figure 7.1. Testing procedure.

Figure 7.2. Y-Balance Test™ directions; a) anterior reach direction; b) posteromedial reach direction; c) posterolateral reach direction.

Figure 7.3. Isometric hip adduction (a) and abduction (b) strength assessment.

Figure 7.4. Lower extremity joint ranges of motion assessment: a) passive hip flexion with knee flexed test [PHF_{KF}]; b) passive hip flexion with knee extended test [PHF_{KE}]; c) passive knee flexion [PKF]; d) passive hip extension [PHE]; e) passive hip abduction [PHA]; f) passive hip external rotation test [PHER]; g) passive hip internal rotation test [PHIR]; h) ankle dorsiflexion with knee flexed test [ADF_{KF}]; and i) ankle dorsiflexion with knee extended test [ADF_{KE}].

Figure 7.5. Participant performing unstable sitting protocol. Projection providing visual feedback of participants' centre of pressure and a target point moving across a circular path.

Figure 7.6. Isokinetic knee flexion and extension strength assessment.

LIST OF TABLES

LIST OF APPENDIXES

ABSTRACT

Soccer (also known as football) requires players to perform many repeated high intensity movements such as sudden acceleration and deceleration, rapid changes of direction, jumping and landing tasks; as well as many situations in which players are involved in tackling to keep possession of or to win the ball. At professional level, the combination of these high physical demands alongside stress and anxiety caused by the congested match calendar may place players at high risk of injury. In fact, soccer is one of the sports with higher injury incidence rates, all of this despite the substantive effort made by the scientific community and physical trainer practitioners to reduce their number and severity. The inefficacy of the preventive measures applied might be caused, in part, by the limitations present in the scientific literature which hinder: a) the accurate estimation of the most frequent soccer-related injuries; b) the identification of professional athletes at high risk of injury; and c) the design of effective neuromuscular training interventions. Therefore, and based on these limitations, the main objectives of the current doctoral thesis were: 1) to carry out a systematic review and a novel meta-analysis of epidemiological data of injuries in professional male soccer; 2) to describe the lower extremity joint ranges of motion profile in professional soccer players; 3) to analyse and compare the behaviour of some machine learning methods in order to select the best performing injury risk factor model to identify professional athletes at risk of lower extremity muscle injuries and hamstring strains; and 4) to analyse the relationships between several parameters of neuromuscular performance with unilateral dynamic balance in such cohort of athletes. To achieve these objectives, a systematic literature review and meta-analysis, a descriptive study, two prospective cohort studies, and a correlational study were conducted.

The main findings of the first study report that professional male soccer players are exposed to a substantial risk of sustaining injuries, especially during matches (32.9 injuries per 1000 hours of player exposure). In particular, the lower extremity is the most frequently injured part of the body, being the thigh the anatomical region in which injuries occurs more. Likewise, the most common type of injury is muscle/tendon strains. On the other hand, the results of study two show the necessity of prescribing exercises aimed at improving hip flexion with knee extended and ankle dorsiflexion with knee flexed ranges of motion within soccer training routines. In addition, as some bilateral deficits were observed, unilateral training should be considered where appropriate. Studies three and four present two different injury risk factor models (personal, psychological and neuromuscular risk factors) to identify players at high risk of lower extremity muscle injuries (area under the receiver operating characteristic curve = 0.747) and hamstring strains (area under the receiver operating characteristic curve = 0.867), respectively. Both models are generated by the SmooteBoost technique with a cost-sensitive alternating decision tree as base classifiers. Finally, the findings of study five indicate that, although male and female professional soccer players report similar unilateral dynamic balance scores, but different measures of neuromuscular performance seem to have influenced this fundamental ability. Thus, for males, those variables related to movement patterns in the sagittal plane (hip flexion and ankle dorsiflexion range of motion measures) were important in the overall balance score obtained. However, for females, variables related to the performance of movement patterns in the frontal plane (such as core stability and hip abduction strength and range of motion) were considered predictor variables of this ability.

Overall, both the results and methodology used in the present doctoral thesis might be used by coaches, physical trainers and clinicians to improve the decision-making process to reduce the number and impact of injuries in professional soccer.

Keywords: *football, injury, prevention, hamstring strain, muscle injury, learning algorithm, data mining, dynamic balance, core stability, performance, range of motion.*

RESUMEN

El fútbol requiere que sus practicantes lleven a cabo un gran número de movimientos repetidos y de alta intensidad, tales como aceleraciones y desaceleraciones súbitas, rápidos cambios de dirección, saltos y caídas; así como muchas situaciones en las que los jugadores están involucrados en luchas por mantener o ganar el balón. A nivel profesional, la combinación de estas altas demandas físicas junto al estrés y ansiedad generada por el intenso calendario competitivo pueden colocar a los jugadores en una situación de alto riesgo de lesión. De hecho, el fútbol es uno de los deportes de equipo con mayores tasas de incidencia de lesiones, todo ello a pesar del sustantivo esfuerzo que en los últimos años ha realizado la comunidad científica y los profesionales de la preparación física para tratar de reducir el número e impacto de éstas. La ineficacia de las medidas preventivas aplicadas hasta la fecha podría deberse, en parte, a las limitaciones que existen en la literatura científica y que dificultan: a) la estimación precisa de las lesiones más frecuentes en el fútbol; b) la identificación de deportistas profesionales en situación de alto riesgo de lesión; y c) el diseño de programas de intervención neuromuscular efectivos. Por lo tanto, y en base a estas limitaciones, los objetivos principales de esta tesis doctoral fueron: 1) llevar a cabo una revisión sistemática y un meta-análisis inédito sobre epidemiología de lesiones en fútbol profesional masculino; 2) describir el perfil de rango de movimiento de las principales articulaciones de la extremidad inferior en jugadores de fútbol profesional; 3) analizar y comparar la habilidad predictiva de un número importante de algoritmos de aprendizaje, fundamentados en árboles de decisión, con el fin de seleccionar el mejor modelo basado en factores de riesgo para identificar jugadores en situación de alto riesgo de lesión muscular de la extremidad inferior o de isquiosural; y 4) analizar las relaciones entre determinados parámetros del rendimiento neuromuscular con el equilibrio dinámico unipodal en dicha muestra de deportistas. Con el propósito de conseguir estos objetivos, se realizó un estudio meta-analítico, dos estudios prospectivos de cohortes, un estudio descriptivo y un estudio correlacional.

Los principales hallazgos del primer estudio informan de que los jugadores profesionales de fútbol están expuestos a un riesgo alto de sufrir lesiones, especialmente durante los partidos (32.9 lesiones por cada 1000 horas de exposición). En particular, la extremidad inferior es la más lesionada, siendo el muslo la región anatómica donde más lesiones se producen. Asimismo, el tipo de lesión más frecuente es la musculo/tendinosa. Por su parte, los resultados del estudio dos muestran la necesidad de prescribir ejercicios destinados a mejorar el rango de movimiento de la flexión de la cadera con rodilla extendida y la flexión dorsal del tobillo con rodilla flexionada durante las sesiones de entrenamiento de fútbol. Además, y dado que los desequilibrios bilaterales del rango de movimiento son frecuentes, el entrenamiento unilateral debería implementarse en caso de ser necesario. Los estudios tres y cuatro presentan dos modelos basados en factores de riesgo (personales, psicológicos y neuromusculares) que permiten identificar a jugadores en situación de alto riesgo de lesión muscular de la extremidad inferior (área bajo la curva característica operativa del receptor = 0.747) y de la musculatura isquiosural (área bajo la curva característica operativa del receptor = 0.867), respectivamente. Ambos modelos están dirigidos por el algoritmo de decisión de árbol alternativo adaptado a un coste específico a favor de la clase minoritaria y modelado en un proceso de ensamblaje (SmooteBoost). Finalmente, los hallazgos del estudio número cinco indican que a pesar de que los hombres y mujeres futbolistas obtienen similares resultados de equilibrio dinámico unipodal, diferentes parámetros del rendimiento neuromuscular parecen influir en esta habilidad. Así, y para los hombres, las variables relacionadas con patrones de movimiento en el plano sagital (rango de movimiento de la flexión de cadera y tobillo) fueron importantes para el equilibrio dinámico. Sin embargo, y para las mujeres, las variables relacionadas con patrones de movimiento en el plano frontal (estabilidad del tronco, fuerza y rango de movimiento de la abducción de cadera) fueron consideradas variables predictoras de dicha habilidad.

De este modo, tanto los resultados como la metodología aplicada en la presente tesis doctoral podrían ser utilizados por entrenadores, preparadores físicos y médicos especialistas para mejorar el proceso de toma de decisiones a la hora de reducir el número e impacto de las lesiones en el fútbol profesional.

Palabras clave: *lesión, prevención, desgarro isquiosurales, lesión muscular, algoritmos de aprendizaje, minería de datos, estabilidad dinámica, estabilidad del tronco, rendimiento, rango de movimiento.*

ABBREVIATIONS

ABD: Abduction.

ACL: Anterior cruciate ligament.

ADD: Adduction.

ADF$_{KE}$: Ankle dorsiflexion with knee extended.

ADF$_{KF}$: Ankle dorsiflexion with knee flexed.

ADTree: Alternating decision tree.

AmC: American cup.

APT: Angle of peak torque.

AT: Artificial turf.

AUC: Area under the receiver operating characteristic curve.

Bila: Bilateral.

BMI: Body mass index.

CI: Confidence intervals.

CON: Concentric.

CoP: Centre of pressure.

CS: Core stability.

CS$_{AP}$: Unstable sitting while performing anterior-posterior displacements with feedback.

CS$_{CD}$: Unstable sitting while performing circular displacements with feedback.

CS$_{ML}$: Unstable sitting while performing medial-lateral displacements with feedback.

CS$_{NF}$: Unstable sitting without feedback.

CS$_{WF}$: Unstable sitting with feedback.

EC: European cup.

ECC: Eccentric.

FUNC: Functional.

H: Hamstring.

HABD: Hip abduction.

HADD: Hip adduction.

HSI: Hamstring strain injury.

ISOK: Isokinetic.

ISOM: Isometric.

KE: Knee extension.

KF: Knee flexion.

kg: Kilograms.

m: Metres.

mm: Millimetres.

MRE: Mean radial error.

MUS_{INJ}: Muscle injuries.

N: Newtons.

NT: National team.

OG: Olympic games.

PHA: Passive hip abduction.

PHE: Passive hip extension.

PHER: Passive hip external rotation.

PHF_{KE}: Passive hip flexion with knee extended.

PHF_{KF}: Passive hip flexion with knee flexed.

PHIR: Passive hip internal rotation.

PKF: Passive knee flexion.

PT: Peak torque.

Q: Quadriceps.

R^2: Explained variance.

ROC: Receiver operating characteristic.

ROM: Range of motion.

ROS: Random oversampling.

RUS: Random undersampling.

SCV: Stratified cross validation.

SD: Standard deviation.

SMOTE: Synthetic minority oversampling technique.

SMT: Smote.

STROBE: Strengthening the reporting of observational studies in epidemiology scale.

TNrate: True negative rate.

TPrate: True positive rate.

UDB: Unilateral dynamic balance.

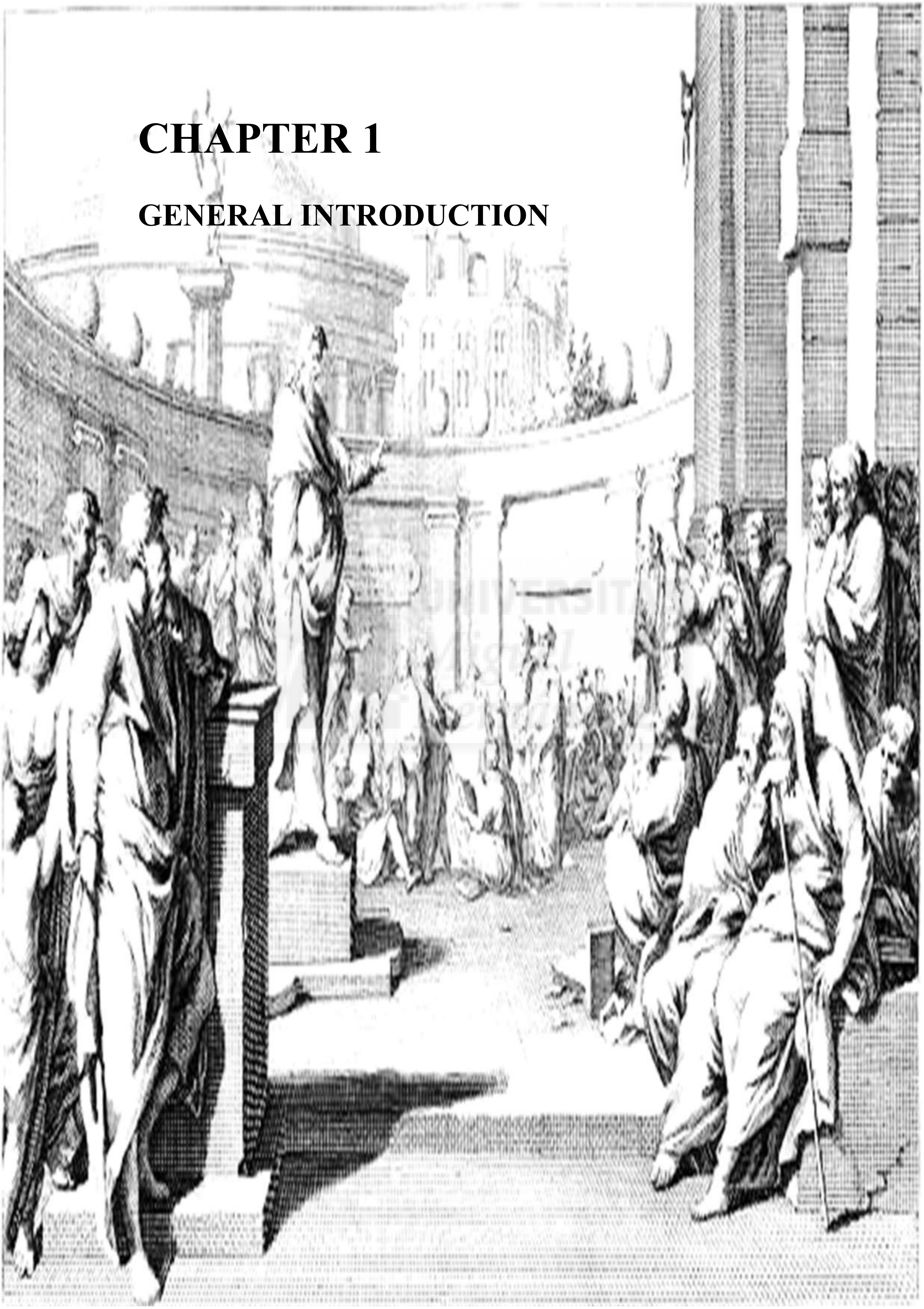
Uni: Unilateral.

WC: Worl cup.

Y-BT: Y-Balance test.

CHAPTER 1

GENERAL INTRODUCTION



CHAPTER 1

GENERAL INTRODUCTION

Soccer (football) is by far the world's most popular sport, with more than 270 million players around the world.[1] At professional level, soccer is clearly oriented to win and achieve the greatest economic benefit. Therefore, soccer health care professionals (among others collectives) work daily (within the ethical limits of their respective profession) to help players and clubs to win.[2] One of the primary strategies to achieve these objectives may be to reduce the number and severity of injuries as it has been demonstrated that their high incidence has a negative impact on a team´s chances of success (e.g. ranking position, games won, goals scored, total points).[3-5]

The way to reduce injury incidence and severity in soccer is called prevention. According to the van Mechelen model for injury prevention,[6] establishing the extent of the injury problem (also called epidemiology) in a predetermined sport population is the fundamental first step to implement adequate injury prevention measures (Figure 1.1). In this sense, soccer health care professionals should know the answers, among others, of the following questions if they want to increase the chances of winning: What are the most frequent injuries? When do they happen? What is the mechanism of these injuries?.

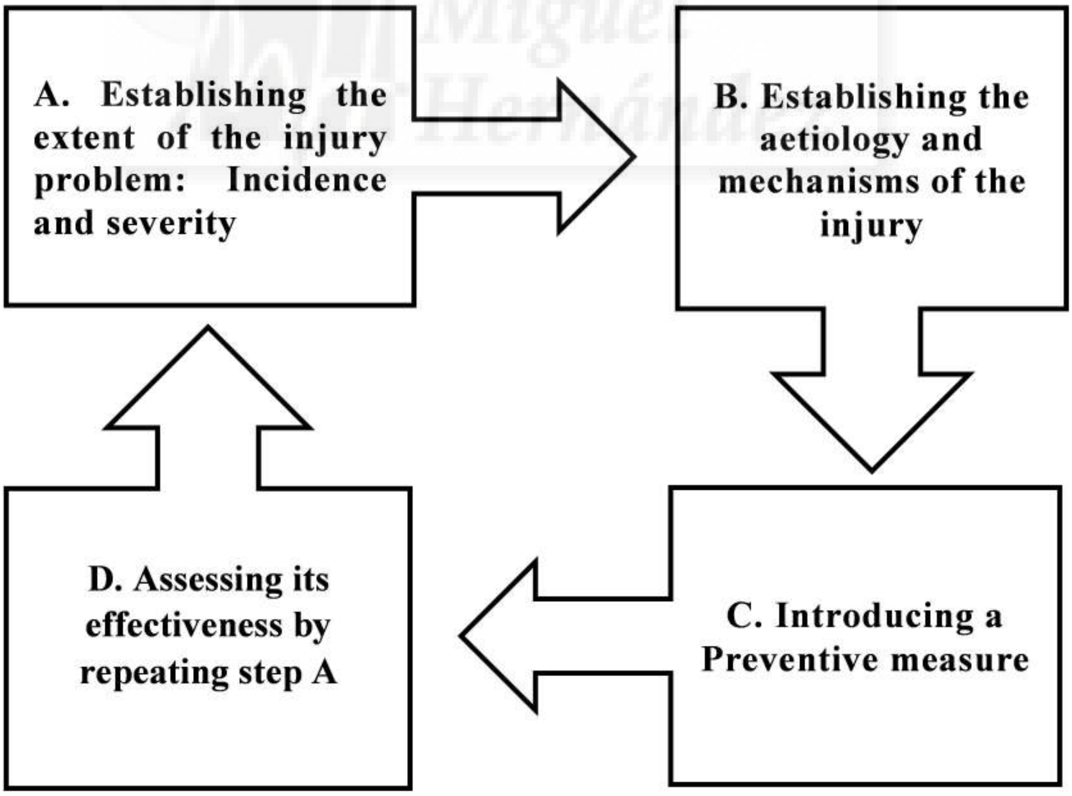


Figure 1.1. The four-step ''sequence of prevention'' described by van Mechelen (1992).

Once the epidemiology (i.e. incidence and severity) of injuries has been carefully analysed, the second step for injury prevention is to establish the aetiology and mechanisms of the injury. This second step lead soccer health care professionals to understand in a better way why injuries happen and to predict who is at risk of injury. To successfully address this step, it is necessary to understand, from a multidimensional and dynamic perspective, the risk factors (and their interactions) and the injury mechanisms that play an important part in the occurrence of the most common injuries.[7-9] This knowledge helps to develop screening programs to identify players at high risk of injury and to apply, when needed, targeted injury prevention measures (third step of van Mechelen model). [6,7]

The following step of the van Mechelen sequence of prevention consists in assessing the effectiveness of the preventive measures implemented by repeating the first step. However, in 2006, Finch[10] redesigned these four steps. She brought to notice the possible gap between the proposed interventions suggested by scientific research and their actual implementation in ‘‘real-life’’ situations. Thus, Finch[10] included modifications that allow for the assessment and application of functional interventions and the determination of the factors that influence safety behaviour in the sporting context. Two years later, Van Tiggelen, Wickes, Stevens, Roosen, & Witvrouw[11] complemented the modifications proposed by Finch by incorporating risk-taking behaviour and compliance of the individual as limiting factors in sports injury prevention.

Therefore, the model of injury prevention initially described by Van Mechelen, Hlobil, & Kemper[6] in 1992 and later expanded by Finch[10] in 2006 and by Van Tiggelen et al.,[11] in 2008 comprises seven steps (Figure 1.2).

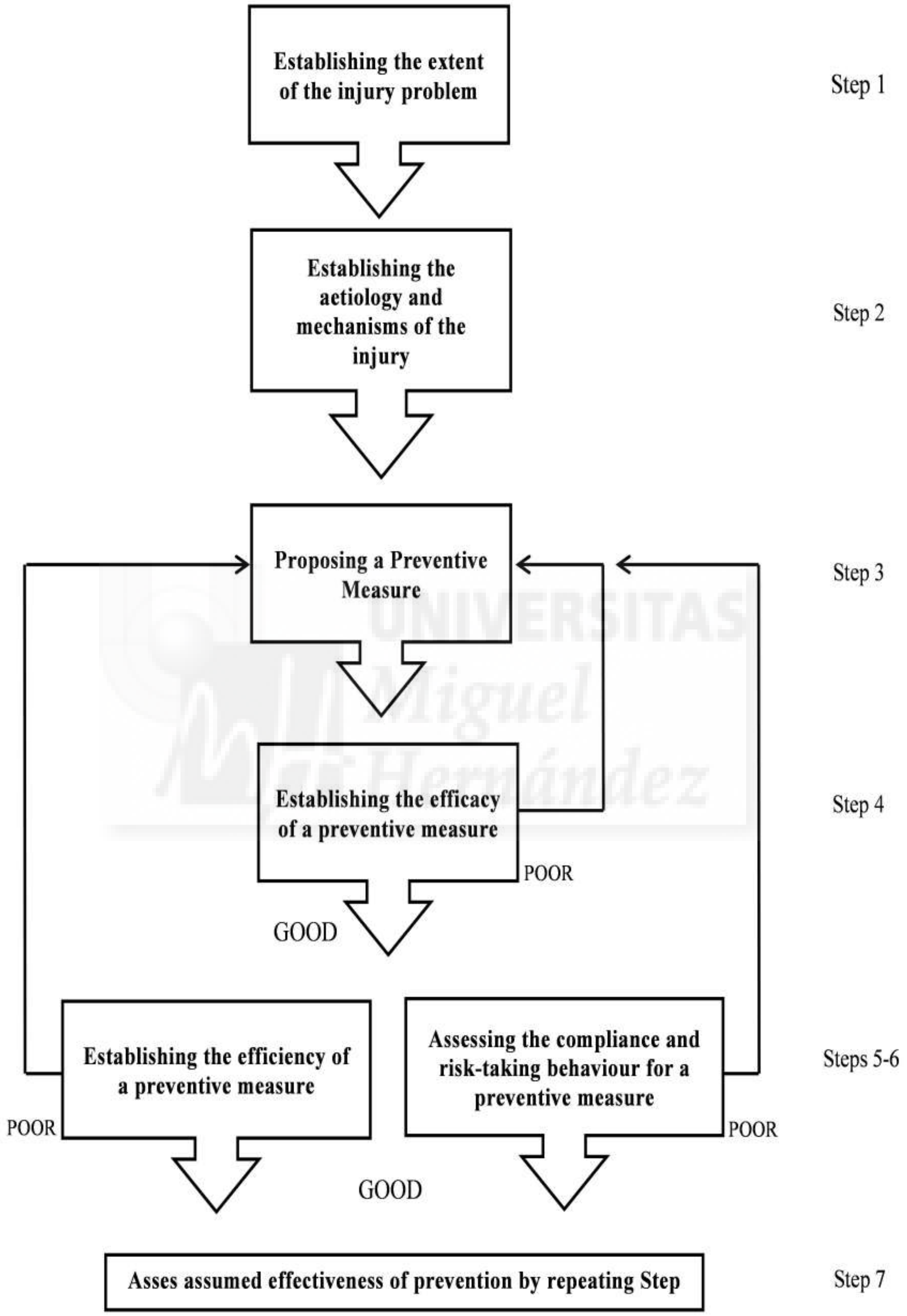


Figure 1.2. Sequence of step of the model of injury prevention described by van Mechelen in 1992 and later expanded by Finch in 2006 and by van Tiggelen in 2008. Figure adapted from Van Tiggelen et al. (2008).

The present doctoral thesis addresses steps number 1, 2 and 3 of the seven-step injury prevention model just presented in figure 1.2 and focuses on professional male soccer players.

1.1. Epidemiology of injuries in professional male soccer

As outlined in the above mentioned seven-step injury prevention model (Figure 1.2), before implementing any preventive measure, the first step is to understand the extent of the problem in terms of the incidence and severity of the injuries.[6,10,11] Therefore, the aim of this section is to briefly summarize the body of the literature regarding the epidemiology of injuries in professional male soccer.

Soccer requires players to perform many repeated high intensity movements such as sudden acceleration and deceleration, rapid changes of direction, jumping and landing tasks, as well as many situations in which players are involved in tackling to keep possession of or to win the ball.[12,13] At professional levels, the combination of these high physical demands[14] alongside situations in which players come into contact (i.e. with the result of a bump, or a kick…), technical evolution[15] and current congested competitive calendars[16,17] may place players at high risk of injury.

In this sense, there have been a number of prospective cohort studies investigating the injuries sustained in soccer players since the end of the 1970s (particularly in countries of the northern hemisphere),[18] and the publication of a consensus statement on injury definitions and data collection procedures in 2006[19] appears to have improved the consistency and quality of research within the field. Thus, in the last two decades several epidemiological studies have been published describing injury patterns over one[16,20-22] or numerous[18,23-25] seasons and during international championships with national teams[26-29] in professional male soccer players. These epidemiology studies have reported that soccer presents one of the highest reported incidence of injuries among the most popular team sports (e.g. basketball, hockey, rugby).[30-32] In particular, the incidence of injuries in professional male soccer ranges from 2.1 to 19.2 injuries per 1000 hours of exposure, being much higher in matches (from 13 to 78.3 injuries per 1000 hours of match exposure) than in training (from 1.5 to 11.8 injuries per 1000 hours of training exposure).[18,33-41] This incidence is much more pronounced in tournament matches with a national team, and can reach up to 101 injuries per 1000 hours of match exposure.[37,42]

Once the extent of the problem has been defined in terms of incidence, the next step is to establish the severity and pattern of the injury. In this sense, recent studies,[18,35,40,43] indicate that most of the injuries have a traumatic origin (injury resulting from a specific, identifiable event) compared to a lower percentage that are caused by overuse (injury caused by repeated micro-trauma without a single, identifiable event responsible for the injury). Overuse injuries are diagnosed mainly in training situations, while traumatic injuries happen especially in match actions.[18]

Although being frequent, traumatic injuries present absence periods of less than one week in 50% of cases, and with 11-16% of them being considered severe and therefore imply more than 28 days of sports absence. In addition, these studies corroborate that the most common injuries are muscle/tendon injuries followed by joint/ligament injuries. On the other hand, the main location of the injuries is the lower extremity with between 70-93% of all injuries,[44] affecting mainly thigh, followed by knee, hip/groin and ankle. In particular, the hamstring strains are those that present a a higher incidence, recurrence and time of absence by injury, with an incremental tendency in recent years.[41,45-48]

However, and despite the fact that a large number of studies reporting data regarding the incidence and severity of soccer-related injuries are available,[18,20,22,37,49,50] no studies (to the author´s knowledge) have combined and meta-analysed such epidemiological data in order to show more robust effect estimates and an increased statistical power. Identifying the most common and severe injuries as well as where (anatomical location) and when (matches or training sessions) they usually occur would lead coaches and physical trainers to prioritise the application of specific measures to prevent or reduce the risk of sustaining such injuries.

Therefore, and despite the aforementioned scientific limitations, there is a clear necessity to develop and implement strategies aimed at preventing and reducing the number and severity of injuries in professional soccer players. However, and prior to establishing injury prevention programmes, it is essential to establish the aetiology and mechanism of the injuries in order to be able to identify soccer players at high risk of injuries through a validated screening programme.[7]

1.2. Prediction of injury risk

Injury prediction is one of the most challenging issues in sport in general and in soccer in particular, and a key component for injury prevention, since the successful identification of injury predictors forms the basis for effective preventive measures.[6] Thus, within the second step of the seven-step prevention model, soccer health care professionals should conduct tests to identify risk factors and players at high risk of injury before proposing preventive measures.

Identifying players at high risk of injury requires the development of a valid and reliable screening program. Screening is a strategy used in a population to detect a disease in individuals without signs or symptoms of that disease. The intention is to identify pathological conditions in early stages, thus enabling an earlier intervention and management in the hope of reducing future morbidity and mortality.[7] Detection of the risk of injury in the sports field usually involves using a performance test to identify certain deficits or imbalances that predispose the individual to injury (range of motion, concentric and eccentric strength, stability, etc.). However, unlike the medical field in which a dichotomous diagnosis is obtained (disease - yes/no), the values obtained in the performance tests are continuous and therefore must be translated into a dichotomous result, that is to say, if the player has a low or high risk of injury.

On this wise, Bahr[7] in a recently published thought-provoking critical review, suggested that prior to considering a screening program as valid to predict and prevent sport injuries it should have overcome three steps successfully (Figure 1.3). The first step is to identify those potential risk factors that have demonstrated a strong relationship with injury in prospective studies and then define appropriate cut-off values. The second step is to determine the validity of the screening tests used to measure the risk factors to predict new injuries in a new athlete population. Finally, in the third step studies should document that an intervention programme targeting athletes identified as being at high risk, using the developed screen, must be more beneficial than the same intervention programme given to all athletes.

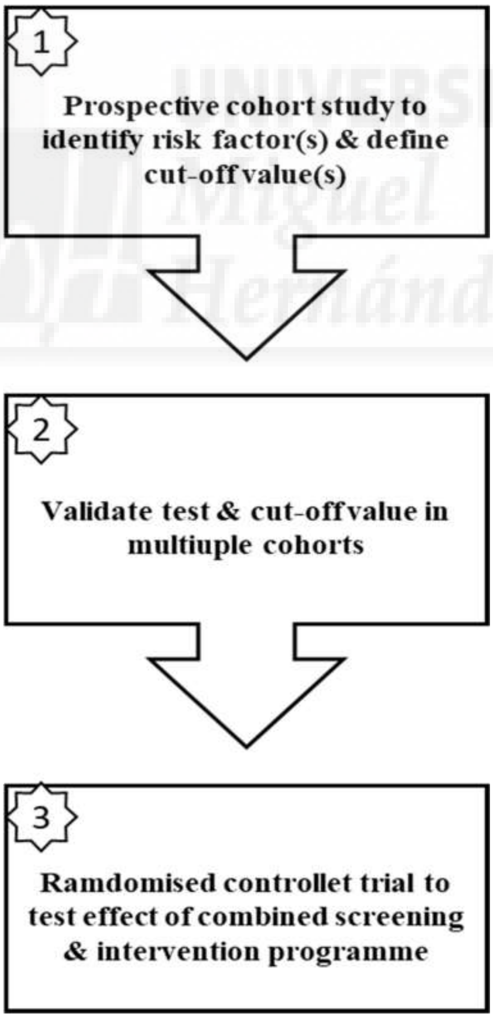


Figure 1.3. The three research steps suggested by Bahr (2016) to develop and validate a screening programme.

In recent years, a substantive effort has been made by the scientific community and medical practitioners to identify strong internal (player-related) and external (environmental-related) risk factors associated with the occurrence of injuries. Thus, some prospective studies, although not all, have identified a number of potential risk factors associated to injury (mainly for muscle/tendon injuries) (Table 1.1).

However, the relationships of certain injury risk factors with injuries are still unclear for some reasons. An example of this circumstance is the classic debate regarding whether the high intensity demands of movements required in soccer could lead to an overload in the joints, generating sport-specific adaptations or causing impairments in players' normal range of motion during football activities and thus may result in a high risk of injuries. In this sense, some studies have analysed the impact of soccer play in some hip (flexion, extension and abduction) and knee (flexion and extension) range of motion measures,[20,51-56] reporting normal (compared to the sedentary population) and non-pathologic (based on the previously published cut-off scores to classify athletes at high risk of injury) values. These results have led some soccer health care professionals to overlook the assessment of the lower extremity joint ranges of motion in preseason screening sessions. However, when interpreting the existing literature regarding the effects of soccer play on normal lower extremity joint ranges of motion, a large degree of inter-player variability in joint ranges of motion is found,[20,51-56] and thus reporting group average range of motion scores may distort the true extent of the number of players reporting restricted range of motion. This aspect alongside other methodological limitations noted, should be addressed before excluding the joint ranges of motion assessment from the preseason screening sessions.

Table 1.1. Injury risk factors.

Internal	**External**
Previous injury[20,23,57,58]	Seasonal distribution[59-62]
Inadequate rehabilitation[20,63,64]	Match associated variables[65-68]
Genetic differences[69,70]	High match and training load[5,17,71,72]
Psychological factors[73,74]	Climatic factors[75-77]
Anthropometric factors[20,72]	Surface[42,50,78]
Physical fitness[79,80]	Warm-up[81,82]
Age[20,75]	Team success[4,66,68]
Sex[45,83,84]	Equipment[62,85,86]
Playing position[34,49,66,75,76,87-89]	
Limb dominance[24,90,91]	
Range of motion[51,56,62,92,93]	
Misalignment[94-96]	
Joint laxity/instability[97,98]	
Level of play[16,22,28,99-104]	
Fatigue[25,28,41,76,105-107]	
Sleep quality[108,109]	

Despite the fact that significant associations (causal relationship) were found between some risk factors and injuries, the ability of the cut-off scores proposed to predict injuries are not acceptable for screening purposes. In particular, most of the cut-off scores reported in previous studies show good true negative rates (i.e. how many individuals with a negative score were not injured); however the true positive rates were very low (i.e. how many individuals with a positive score were injured). Thus, and for example Hewett et al.,[110] introduced the vertical drop jump test as a screening test for anterior cruciate ligament injury in female athletes and they observed the strongest association with injury risk of peak external knee abduction moment during landing, concluding that this factor predicted anterior cruciate ligament injury status with 78% sensitivity and 73% specificity. However, other groups have examined the vertical drop jump test and they have not been able to confirm that there is an association between knee abduction and injury risk.[111,112]

Several arguments seem to be behind the lack of generality of the proposed cut-off values and it could explain why they do not allow the identification of athletes at high risk of injury. Perhaps one the main reasons behind the lack of available valid screening programmes to predict athletes at high risk of suffering a sport injury could be based on the use of traditional and/or sub-optimal statistical approaches (e.g. multivariate logistic regression analysis). Logistic regression models generally do not deal well with class imbalance problems, such as the injury phenomenon, in which the number of injured players (minority class) prospectively reported is always much lower than the non-injured players (majority class).[113] Thus, in many scenarios including injuries, traditional multivariate analyses are often biased (for many reasons) towards the majority class (known as the ‘‘negative’’ class) and therefore, there is a higher misclassification rate for the minority class instances (called the ‘‘positive’’ examples), which represent the most important concept.[114]

Furthermore, another limitation of the current body of the literature is based on the fact that most of the predictive models available in the sports medicine context might show a limited external validity. In this sense, the generality (external validity) of the cut-off scores proposed for certain injury risk factors (e.g. strength imbalances, joint ranges of motion) might be limited since their predictive abilities to identify new athletes at high risk of injury have not been verified in a new population of athletes, different from the one used for defining them (i.e. cross validation).[7,115] This suggests that cut-off scores might be overfitted (i.e. their predictive ability is adjusted to the data set used in their learning process), which will give overly optimistic results and hence, they may not be acceptable for screening purposes. This appears to be supported by the fact that the cut-off scores defined by some prospective studies (mainly those related to strength measures) have not been later ratified by others studies using similar designs and assessment methodologies but with different samples of athletes.[20,23,93,116-125] For example, while Croisier, Ganteaume, Binet, Genty, & Ferret[118] and Dauty, Menu, Fouasson-Chailloux, Ferréol, & Dubois[120] found that professional soccer players with reciprocal (functional) hamstring-to-quadriceps ratios lower than 0.8 were at higher risk of sustaining a hamstring strain, van Dyk et al.,[125] did not identify this strength ratio measure as a risk of hamstring strain injury.

Finally, another issue with the current evidence base is that the above-mentioned prospective studies have identified potential risk factors for injury according to the presence of statistically significant relationships (based on odds ratios, certain values of p statistic [mainly $p < 0.05$]) with injuries. However, based on the general agreement that the aetiology of injury is multifactorial and that some relationships of conditional dependence might exist among factors,[126] it is possible that the influence of a specific factor on the likelihood of suffering an injury might not be statistically significant ($p < 0.05$) in itself, but relevant when it is used in conjunction with several other factors to develop a more robust predictive model. In other words, combining information from several internal and external risk factors might lead to the development of a more robust model with an improved predictive ability.

The consequence of this has led Bahr[7] to conclude that: a) to find a statistically significant association between a predictive test result and an injury is not sufficient evidence to use the test to predict who is at risk of injury; and b) there is no screening test available to predict sport injuries with adequate properties and consequently the exercises included in intervention programmes are not evidence-based supported, as the link between risk factors and injury incidence remains to be established.

The application of mechanical learning techniques might solve the extent injury problem in professional soccer, in which a large number of factors are involved and the use of resampling techniques (e.g. cross-validation, bootstrap and leave-one-out) may overcome the limitations inherent to the current body of knowledge and give light to better identification of athletes at high risk of injury.[127]

For example, statistical methods of data mining and mechanical learning have already been used in sport performance analysis,[128-132] while to date (to the author´s knowledge), only one study has used these statistical systems for injury prediction.[133]

In particular, decision tree algorithms (Figure 1.4) are powerful statistical tools for prediction that have been used in several medical diagnosis studies reporting excellent results.[134-137] These learning algorithms have several advantages compared to traditional approaches (e.g. logistic regression)[138]: a) they simplify complex relationships between input variables and target variables by dividing the original input variables into significant subgroup; b) they are easy to understand and interpret; c) they use non-parametric approach without distributional assumptions; d) it is easy to handle missing values without needing to resort to imputation; e) it is easy to handle heavy skewed data without needing to resort to data transformation; and f) they are robust to outliers.

In addition, decision tree algorithms are also able to select only those variables that are considered necessary to develop poweful predictive models, reducing the number of variables necessary for a decision-making process. They can also be used as based classifiers in ensemble methods (also known as multiple classifier systems) to obtain better predictive performance scores.[139] Finally, the implementation of resampling techniques (e.g. cross validation, which involves [repeated multiple times] that a subset of the sample [training subset] is used to fit a model and the remaining subset [test subset] is used to estimate the efficacy of the model) to validate the model might provide a more accurate estimation of the predictive performance and increase its generality.[115]

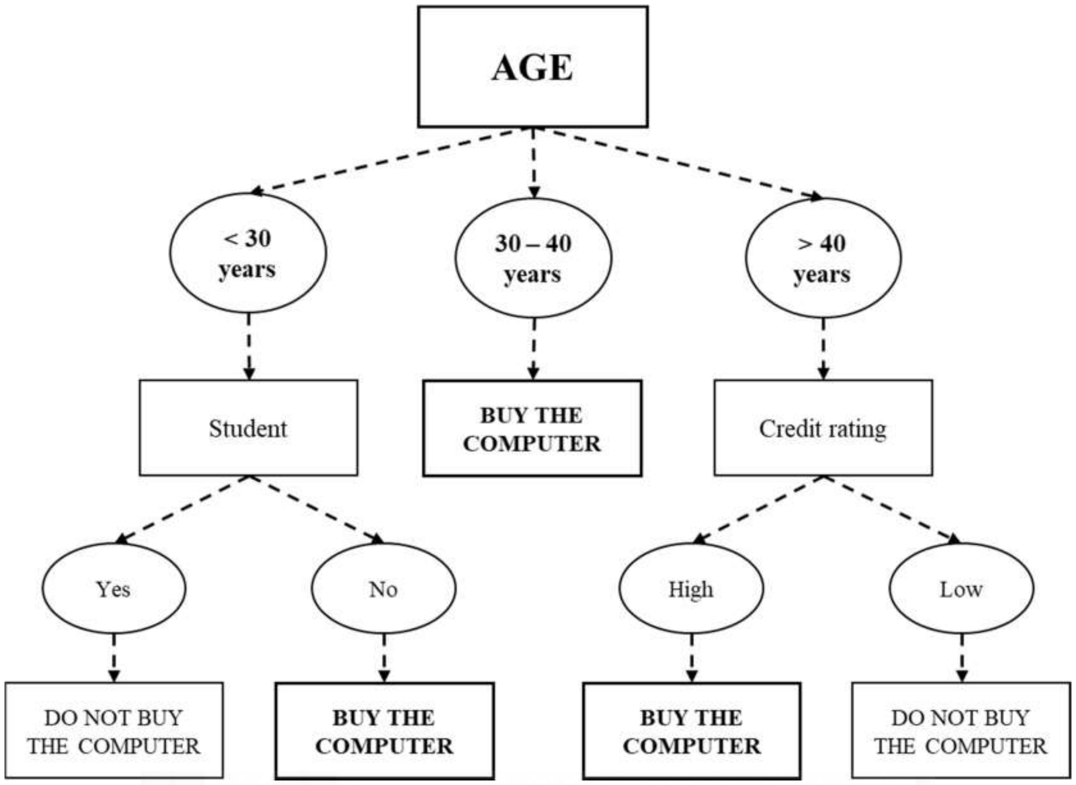


Figure 1.4. Example of a simple decision tree for buying a computer.

1.3. Proposing preventive measures

Once soccer health care professionals have identified players at high risk of injury, the following step in the injury prevention model presented is to propose preventive measures designed to correct those impairments found in the musculoskeletal and/or sensorimotor systems.

Deficits or impairments in the unilateral dynamic balance (defined as the ability of an individual to maintain the center of mass within the body´s base whilst performing single leg movements)[140] are frequently reported in soccer players. [97] Therefore, it seems necessary to identify which measures of neuromuscular performance (e.g. hip and knee strength, lower extremity joint ranges of motion, core stability) could have an impact on unilateral dynamic balance in order to design targeted preventive measures.

Although some studies have explored the individual contribution of certain modifiable measures of neuromuscular performance on unilateral dynamic balance in soccer (knee[141-143] and hip[144] strength, jumping ability,[141,145] core stability,[145] ankle dorsiflexion[146,147] and hip flexion[147] ROMs) only one study has used professional players.[141] In addition, to the author´s knowledge, no studies have analysed the concurrent influence of the main training modifiable neuromuscular measures on the unilateral dynamic balance in soccer players. Finally, and although previous studies have reported no sex-related difference in unilateral dynamic balance in college athletes,[148] basketball players[149] and recreational athletes,[150] none of these studies has determined if males and females used similar or different neuromuscular strategies to achieve a better unilateral dynamic balance. Consequently, the relationships between the main training modifiable measures of neuromuscular performance with unilateral dynamic balance in professional soccer players remain unresolved. This knowledge would allow clinicians and sport practitioners to develop more effective and tailored unilateral dynamic balance training programmes in soccer players, possibly improving performance and reducing the risk of injury.

1.4. Lines of action of the thesis

Once the scientific literature behind the first three steps of the injury prevention model has been addressed, the need to reduce the number of injuries in soccer is clear. Likewise, it is necessary to resolve the limitations encountered in the world of sports medicine to deal with this problem, since as it has been presented in this introduction, the numerous attempts to reduce injuries in soccer have not had the desired effects. Therefore, within the first step of injury prevention, it is essential to develop epidemiological studies that explore and analyse all existing data on the extent of injuries in soccer, both in incidence and severity. In a second step, given the lack of efficient predictive models in the prediction of sports injuries, it seems necessary to develop studies with novel and more complex statistical models that allow both to discern between players with high or low risk of injury and to establish potential risk factor on the injury, which could cause specific sports adaptations and consequently make the player more susceptible to an injury or an adaptation of these factors. In the third step, establishing the effect of each risk factor on the injury and possible dynamic interrelationships between them might provide more effective and individualized preventive measures (Figure 1.5). In doing so, novel research questions arise that will be of interest to clinicians, coaches, players and football organizations.

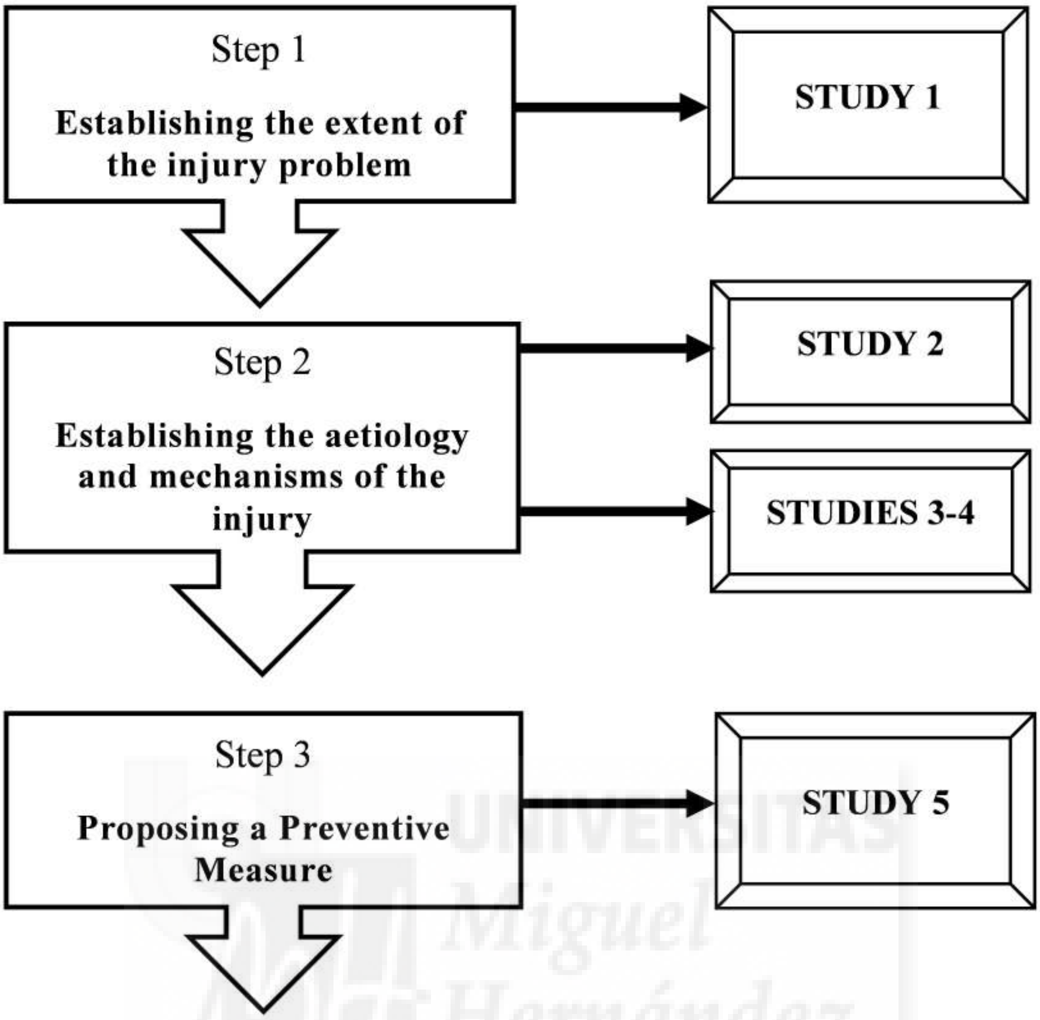


Figure 1.5. Van Tiggelen prevention model with studies included in this thesis.

CHAPTER 2

RESEARCH OBJECTIVES AND HYPOTHESES

CHAPTER 2

RESEARCH OBJECTIVES AND HYPOTHESES

2.1. General objectives

Based on the limitations of the literature, the general objectives of this doctoral thesis were *to establish the extent of the soccer injury problem and its risk factors to develop innovative predictive models for identifying professional soccer players at high or low risk of injury during preseason screening,* and *to deepen the knowledge and relationships between the different risk factors in soccer in order to be able to develop prevention strategies for each individual player.*

To achieve the referred objectives, five studies were performed. The first study described the extent and nature of soccer injuries through an epidemiology systematic review and a novel meta-analysis. The second study described lower extremity range of motion in professional soccer players. The following two studies utilized learning algorithms to develop lower extremity muscle and hamstring strain injury prediction models to identify professional soccer players at high or low risk of injury. It should be clarified that although studies three and four included a limited number of professional handball players (n = 34) as participants, the analysis of the epidemiology of injuries and the relationships existing among neuromuscular risk factors in this cohort of players is out of the scope of the current doctoral thesis. The rationale of recruiting handball players as participants was because of the need to assess possible impact of the sport modality as a variable on the predictive ability of the models generated and because both sports have high levels of injury incidence with similar movement patterns, physiological demands and injury patterns. The last study explored the interaction between several neuromuscular performance parameters and certain sex-related differences.

The titles of the five studies are the following:

- **Study 1**: Epidemiology of injuries in professional soccer injuries: a systematic review and meta-analysis.
- **Study 2**: Comprehensive profile of hip, knee and ankle ranges of motion in professional soccer players.
- **Study 3**: A preventive model for muscle injuries. A novel approach based on learning algorithms.
- **Study 4**: A preventive model for hamstrings muscle injuries in professional soccer and handball players: A novel approach based on learning algorithms.
- **Study 5**: Relationships and sex-related differences in several neuromuscular parameters with dynamic balance in soccer players.

2.2. Specific objectives

The specific objectives have been structured depending on the five studies of this doctoral thesis:

Study 1:

1. To carry out a systematic review and a novel meta-analysis of epidemiological data of injuries in professional male soccer as reported in the literature.
2. To make magnitude-based inferences regarding location of injuries, type of injuries, severity of injuries, overuse vs. traumatic injuries, new vs. recurrent injuries, level of play and trend in injury incidence over time.

Study 2:

3. To describe the lower extremity range of motion profile in professional soccer players.
4. To analyse if there are differences between goalkeepers and outfield players in the lower extremity range of motion using robust methods of measure.

Study 3:

5. To analyse and compare the predictive ability of a wide range of decision tree learning algorithms to select the best performing injury risk factor model to identify professional players at high risk of muscle injuries.

Study 4:

6. To analyse and compare the predictive ability of a wide range of decision tree learning algorithms to select the best performing injury risk factor model to identify professional players at high risk of hamstring strain injuries.

Study 5:

7. To analyse the relationships between several parameters of neuromuscular performance with unilateral dynamic balance measured through a Y-Balance Test.
8. To determine the possible sex-related differences in the overall balance score in a cohort of professional soccer players and if males and females use different neuromuscular strategies to achieve them.

2.3. Research hypotheses

The following hypotheses were established in the five studies of this doctoral thesis:

Study 1:

1. Based on previous studies on soccer epidemiology,[18,20,37] the incidence of match injuries will most likely be higher than the incidence of injuries in training, showing an incremental trend over the years. This will be due to the increase of the matches' physical demands and the congested match calendar.[13,151,152]
2. According to the literature,[17,20,41] muscle/tendon injuries in the lower extremity, particularly the thigh, will be the most frequent injuries.
3. Although most studies have indicated that minor/mild injuries are the most frequent,[18,22,25] our meta-analysis will show slight/minimal injuries as more common due to the improvement of the injury registry system,[19] the preventive programs and the post-exercise recovery techniques.[153]

Study 2:

4. In line with the literature,[51,52,56] soccer will not generate specific adaptations in lower extremity joint ranges of motion or bilateral differences between legs.
5. The different physical and physiological demands between goalkeepers and field players,[154,155] will produce greater range of motion in goalkeepers´ lower extremity than in field players.

Study 3:

6. The application of complex statistical approaches coming from machine learning and data mining environments which have been very successful in sport performance analyses,[128-132] will provide a novel model to predict muscle injuries with a higher external validity than traditional multifactorial approaches.

Study 4:

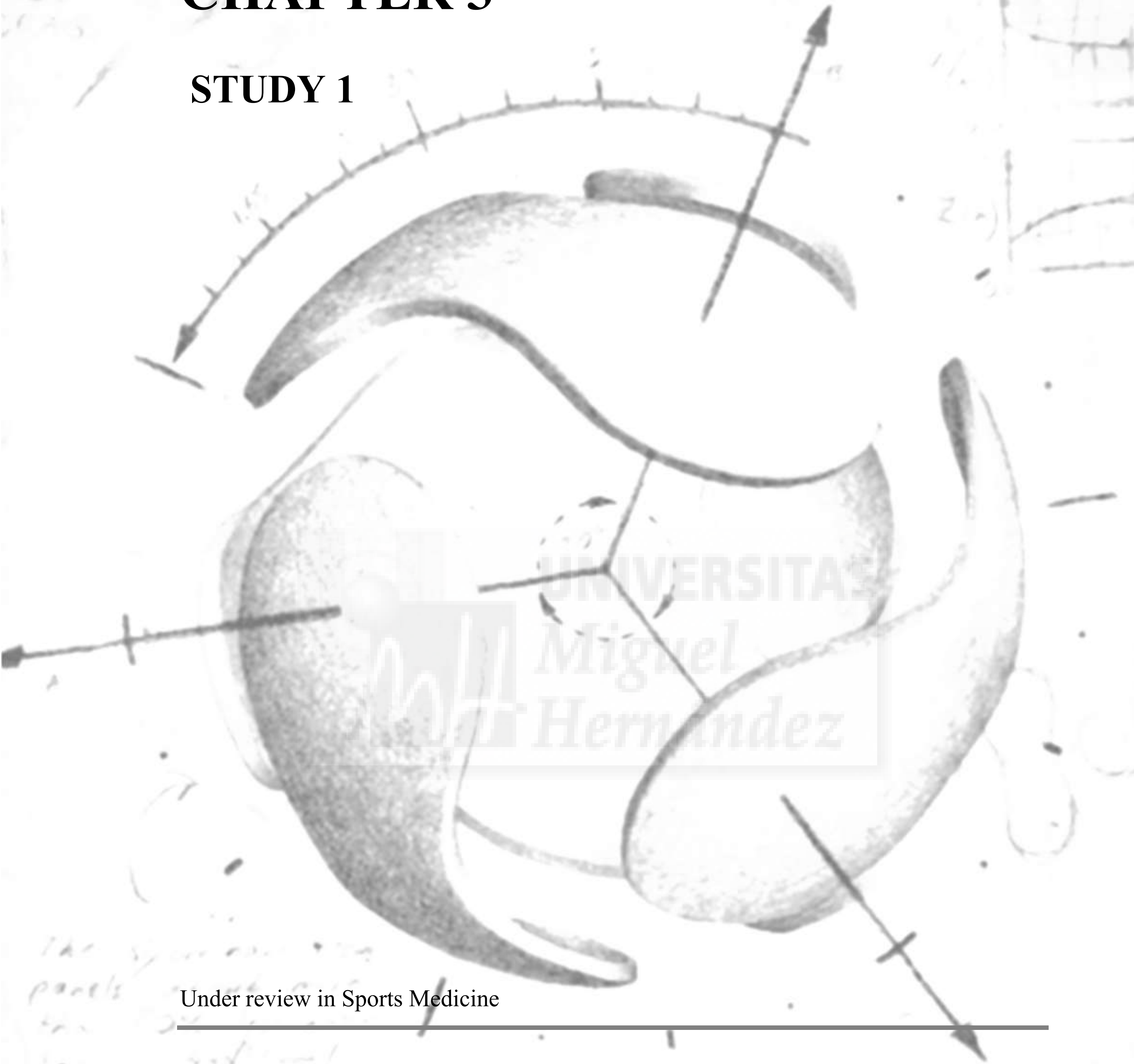
7. The application of complex statistical approaches coming from machine learning and data mining environments will improve current reductionist paradigm on hamstring strain injuries,[126] providing a novel and multidimensional model to predict players at high or low risk of hamstring injury.

Study 5:

8. Based on previous studies on the interaction of different neuromuscular parameters in unilateral dynamic balance,[52,141,144] this study will identify new neuromuscular parameters with high relevance on this ability, measured with the Y-Balance test.
9. There will be no sex-related differences in the Y-Balance test performance, considering previous studies on sex-related differences in unilateral dynamic balance.[149,156]

CHAPTER 3

STUDY 1



Under review in Sports Medicine

Epidemiology of injuries in professional soccer: a systematic review and meta-analysis

Alejandro Lopez-Valenciano, Iñaki Ruiz-Pérez, Jose Alberto García Gómez, Francisco J. Vera-Garcia, Mark De Ste Croix, Gregory Myer, Francisco Ayala Rodríguez.

CHAPTER 3

STUDY 1

Epidemiology of injuries in professional soccer: a systematic review and meta-analysis

by

López-Valenciano A, Ruiz-Pérez I, García-Gómez A, Vera-Garcia FJ, De Ste Croix M, Myer G, & Ayala F.

3.1. Abstract

Background: Despite the fact that a large number of studies have reported data exploring the incidence and severity of soccer-related injuries, no studies have combined and meta-analysed such epidemiological data in order to show more robust effect estimates and an increased statistical power.

Objective: The aim of this study was to carry out a systematic review and meta-analysis of epidemiological data of injuries in professional male soccer.

Method: Forty-nine prospective and three retrospective cohort studies reporting the incidence of injuries in soccer were included. Two reviewers independently extracted data and assessed trial quality using the STROBE statement. Studies were combined in a pooled analysis using a Poisson random effects regression model. Magnitude based inferences were used to assess differences between factors.

Results: The overall incidence of injuries in professional male soccer players was 7.7 injuries/1000 hours of exposure, 3.9 injuries/1000 hours of training exposure and 32.9 injuries/1000 hours of match exposure. Lower extremity injuries had the highest incidence rates compared to other body regions. The most common type of injury was muscle/tendon, which were frequently associated to traumatic incidents. Minimal injuries (1-3 days) were the most usual injuries. A slight upward tendency (non-statistically significant) in match injury incidence over the last three decades was observed.

Conclusions: Professional male soccer players are exposed to a substantial risk of sustaining injuries, especially during matches. In order to markedly reduce overall injury burden, efforts should target lower-extremity injury prevention strategies and movement technique during contact situations, as these may render the largest effect.

Keywords: *football, incidence, lower extremity, muscle/tendon, time-loss.*

3.2. Introduction

Soccer is by far the most popular sport in the world.[1] Soccer requires players to perform a high number of repeated high intensity movements such as sudden acceleration and deceleration, rapid changes of direction, jumping and landing tasks, as well as many situations in which players are involved in tackling to keep possession of or to win the ball.[12,13] At professional levels, the combination of these high physical demands alongside exposure to contacts situations and current congested competitive calendars may place players at high risk of injury. In fact, soccer presents one of the highest reported incidences of injuries among the most popular sports (e.g. basketball, hockey, rugby).[30,157,158] For example, a professional soccer team with a 25 player-squad typically suffers about 50 injuries that cause time-loss (mean of 18 days) from play each season, which equates to two injuries per player per season.[18] The impact of injuries on team performance can therefore be considerable. Indeed, it has been demonstrated that lower injury incidence rates have strong correlations ($r > 0.85$) with team success (e.g. ranking position, games won, goals scored, total points).[4,5] Furthermore, given the financial and competitive concerns in professional soccer, another significant feature of injuries is that their overall burden could be very significant for clubs. In particular, the average cost of a first-team player in a professional team being injured for 1 month is calculated to be around €500.000.[159]

Therefore, there is a clear necessity to develop and implement measures aimed at preventing and reducing the number and severity of injuries in professional soccer players. However, prior to implementing injury prevention programmes, it is essential to establish the extent of the problem in terms of the incidence and severity of injuries[6,10]. There have been a number of prospective cohort studies investigating the injuries sustained in soccer players since the end of the 1970s,[18] and the publication of a consensus statement on injury definitions and data collection procedures in 2006[19] appears to have improved the consistency and quality of research within the field. Thus, in the latter two decades several epidemiological studies have been published describing injury patterns over one[16,20-22] or numerous[18,23-25] seasons and during tournaments with national teams[26-29] in professional male soccer players. However, and despite the fact that a large number of studies reporting data regarding the incidence and severity of soccer-related injuries are available,[18,20,37,49,83] no studies (to the authors´ knowledge) have combined and meta-analysed such epidemiological data in order to show more robust effect estimates and an increased statistical power. Identifying the most common and severe injuries as well as where (anatomical location) and when (match or training sessions) they usually occur would help coaches and physical trainers to prioritise the application of specific measures to prevent or reduce the risk of sustaining such injuries.

Therefore, the main purpose of the current study was to carry out a systematic review and a novel meta-analysis of epidemiological data of injuries in professional male soccer as reported in literature, as well as to make magnitude-based inferences regarding location of injuries, type of injuries, severity of injuries, overuse vs. traumatic injuries, new vs. recurrent injuries, level of play and trend in injury incidence over time.

3.3. Method

To carry out this study, guidelines for reporting meta-analysis of observational studies in epidemiology (PRISMA guidelines) were followed.[160]

3.3.1. Study selection

To be included in the meta-analysis, the studies had to fulfil the following criteria: 1) the study design must be prospective or retrospective; 2) injury had to be defined in term of time-loss;[19,161] 3) participants had to be professional or elite male adult (aged ≥ 18 years) soccer players; 4) the study had to be published in a peer-reviewed journal before June 2017; and 5) it must be written in English or Spanish. Furthermore, 6) eligible studies had to report either incidence rate or prevalence period among the surveyed players, or provide sufficient data from which these figures could be calculated. Studies using injury definitions other than time-loss were excluded. Literature reviews, abstracts, editorial commentaries and letters to the editors were also excluded. Finally, some authors were contacted to provide missing data or to clarify if data were duplicated in others publications. Incomplete data, or data from an already included study, were excluded.

3.3.2. Search strategy

Potential studies were identified by combined search processes, clearly planned and ordered. Firstly, the following bibliographical databases were consulted: PubMed, Scopus, EMBASE, AMED, Google Scholar and the Cochrane Library with the following search terms included in Boolean search strategies: ("soccer" OR "football" NOT “rugby”) AND ("injury" OR "injuries") AND ("professional" OR "elite"). By using filter criteria of the respective databases, the search was limited to publication dates (to 2017/6/31), human species, males, and English and Spanish languages. Secondly, several specialized electronic journals were also consulted, including: American Journal of Sports Medicine, British Medical Journal, Scandinavian Journal of Medicine & Science in Sports, European Journal of Sports Sciences and British Journal of

Sports Medicine. Finally, the reference lists of the studies recovered were hand-searched to identify potentially eligible studies not captured by the electronic searches.

Two reviewers independently (A.L. and I.R.): a) screened the title and abstract of each reference to locate potentially relevant studies, and once hardcopies of the screened documents were obtained; b) reviewed them in detail to identify articles that met the selection criteria. A third external reviewer (F.A.) was consulted to resolve discrepancies.

3.3.3. Data extraction

With the aim of guaranteeing the maximum possible objectivity, a codebook was produced that specified the standards followed in coding each of the characteristics of the studies. The moderator variables of the eligible studies were coded and grouped into three categories: 1) general study descriptors (i.e. authors, year of publication and study design); 2) description of the study population (i.e. sample size, age and level of play); and 3) main epidemiologic findings (i.e. injury and exposure data, distribution of injuries by anatomic location, type of injury and injury severity). If applicable, the authors of included studies were contacted to provide clarifications or access to raw data. The purpose of the current meta-analysis was to determine the overall effects of: 1) soccer-related injury incidence (overall vs. training vs. match injuries rates); 2) location of injuries (lower extremity vs. trunk vs. upper extremity vs. neck vs. head); 3) type of injuries (fractures and bone stress vs. joint [non-bone] and ligament vs. muscle and tendon vs. contusions vs. laceration and skin lesion vs. central/peripheral nervous system vs. undefined/other); 4) severity of injuries (slight/minimal [1-3 days] vs. minor/mild [4-7 days] vs. moderate [8-28 days] vs. major/severe [> 28 days]); 5) mechanism of injury (overuse vs. traumatic injuries); 6) new vs. recurrent injuries; 7) level of play (national level [national leagues] vs. international level [tournaments with national teams]). Thus, multiple rows of data were included for each study to allow for the various combinations of counts and exposures required for each random effect. Additionally, a descriptive analysis was provided to describe trends in injury risk over time. Appendix 3.1 displays a brief description of the moderator variables coded separately by category.

3.3.4. Quality assessment

Two reviewers independently assessed the reporting quality of included studies using an adapted version of the “Strengthening the Reporting of Observational Studies in Epidemiology” (STROBE) statement.[162] All included studies were rated on 11 specific criteria which were derived from items 5, 6, 7, 8, 9, 12, 14 and 15 of the original checklist. The STROBE scale has been considered as a suitable starting point for assessing quality of observational studies.[163] Moreover, the STROBE statement checklist has been adopted as a quality assessment tool by a number of authors.[164,165] This 11-item checklist provides guidance on the reporting of observational studies in order to facilitate critical appraisal and interpretation of results. The observational studies were considered as having a low risk of bias if they were determined as high quality (score of ≥ 7/11) or a high risk of bias if they were low quality (< 6/11).[162] Final study ratings for each reviewer were collated and examined for discrepancies.

To assess the inter-coder reliability of the coding process, two researchers coded 25 studies randomly (48%) (including methodological quality assessment). For the quantitative moderator variables, intra-class correlation coefficients (ICC) were calculated, while for the qualitative moderator variables, Cohen’s kappa coefficients were applied. On average, the ICC was 0.89 (range: 0.78-1.0) and the kappa coefficient was 0.90 (range: 0.81-1.0), which can be considered highly satisfactory, as proposed by Orwin, & Vevea.[166] The inconsistencies between the two coders were resolved by consensus, and when these were due to ambiguity in the coding book, this was corrected. As it has been explained before, any disagreement was resolved by mutual consent in consultation with a third reviewer. The codebook can be obtained from the corresponding author.

3.3.5. Statistical analysis

Injury incidence rates per 1000 hours of player exposures were extracted from the included studies. If injury incidence rates were not specifically reported, they were, if possible, calculated from the available raw data.

Data were modelled by a random effects Poisson regression model, as previously described.[30,167] The response variable was the number of observed injuries, offset by the log of the number of exposure hours. A random effects term was included to account for the correlation arising from using multiple rows of data from the same study. Factors of interest were included as random effects. The weighting factor used was: study exposure time [hours]/mean study exposure time [hours]. The possible influence of the following variables on the model was analysed independently and in conjunction through univariate and multivariate analyses: weeks of follow-up; year of the study publication, age of the players, STROBE score and number of teams included in the study.

Heterogeneity was evaluated using the I^2 statistic, which represents the percentage of total variation across all studies due to between-study heterogeneity.[168] All statistical analyses were performed using the statistical software package R Version 2.4.1 (The R Foundation for Statistical Computing) and the ‘‘metafor’’ package.[169]

For injury incidence data, the overall estimated means for each random-effect factor were obtained from the model and then back-transformed to give incidence rates, along with 95% confidence intervals (CIs). Comparisons between factors were then made using a spreadsheet for combining effect statistics,[170] whereby the incidence rate ratio (and its associated confidence limits) was assessed against predetermined thresholds. An incidence rate ratio of 0.91 represented a substantially lower injury risk, while an incidence rate ratio of 1.10 indicated a substantially higher injury risk.[171] An effect was deemed unclear if its confidence interval overlapped the thresholds for substantiveness; that is, if the effect could be substantial in both a positive and negative sense. Otherwise the effect was clear and deemed to have the magnitude of the largest observed likelihood value. This was qualified with a probabilistic term using the following scale[172,173]: <0.5%, most unlikely; 0.5–5%, very unlikely; 5–25%, unlikely; 25–75%, possible; 75–95%, likely; 95–99.5%, very likely; > 99.5%, most likely.

3.4. Results

3.4.1. Descriptive characteristics of the studies

A total of 1687 references were identified with all search strategies, from which 1261 were excluded in the first screening as duplicates (approximately 75%). Three hundreds and eight studies (approximately 18%) were eliminated after reading the title and abstract. Another 20 studies did not report injury incidence rates or were focused on specific types of injuries (e.g. ankle sprains, hamstrings muscle strains) (about 1%). Forty-six did not methodologically comply with the established criteria such as injury definition, participants observed (amateur players, female or children) (close to 2%) and data duplication.

The search process identified 52 articles (resulting in 69 cohort groups as 12 studies had more than one group) that met the inclusion criteria.[16,18,20-24,26-29,33-43,49,50,71,77,83,88,99,100,106,174-194] Forty-nine articles were prospective cohort studies and three were retrospective cohort studies.[71,175,193] Figure 3.1 shows the flow chart of the selection process of the studies. The studies were carried out between 1989 and 2016 and comprised 78300 injuries among 32700 players from both tournaments with national teams (world[26-28,106,175,177,192] and continental[29,37,193] cups) and professional soccer leagues in a large number of countries[18,22,35,77,182] (e.g. United Kingdom, Denmark, Spain, Sweden, Norway, United States, Iceland, Italy, France). The mean age of the players was 24.6 ± 2.2 years. The number of teams comprised of more than 1200 with more than 7492 weeks of follow-up, which equates to approximately 195 seasons. Table 3.1 provides a descriptive summary of the characteristics of the included studies.

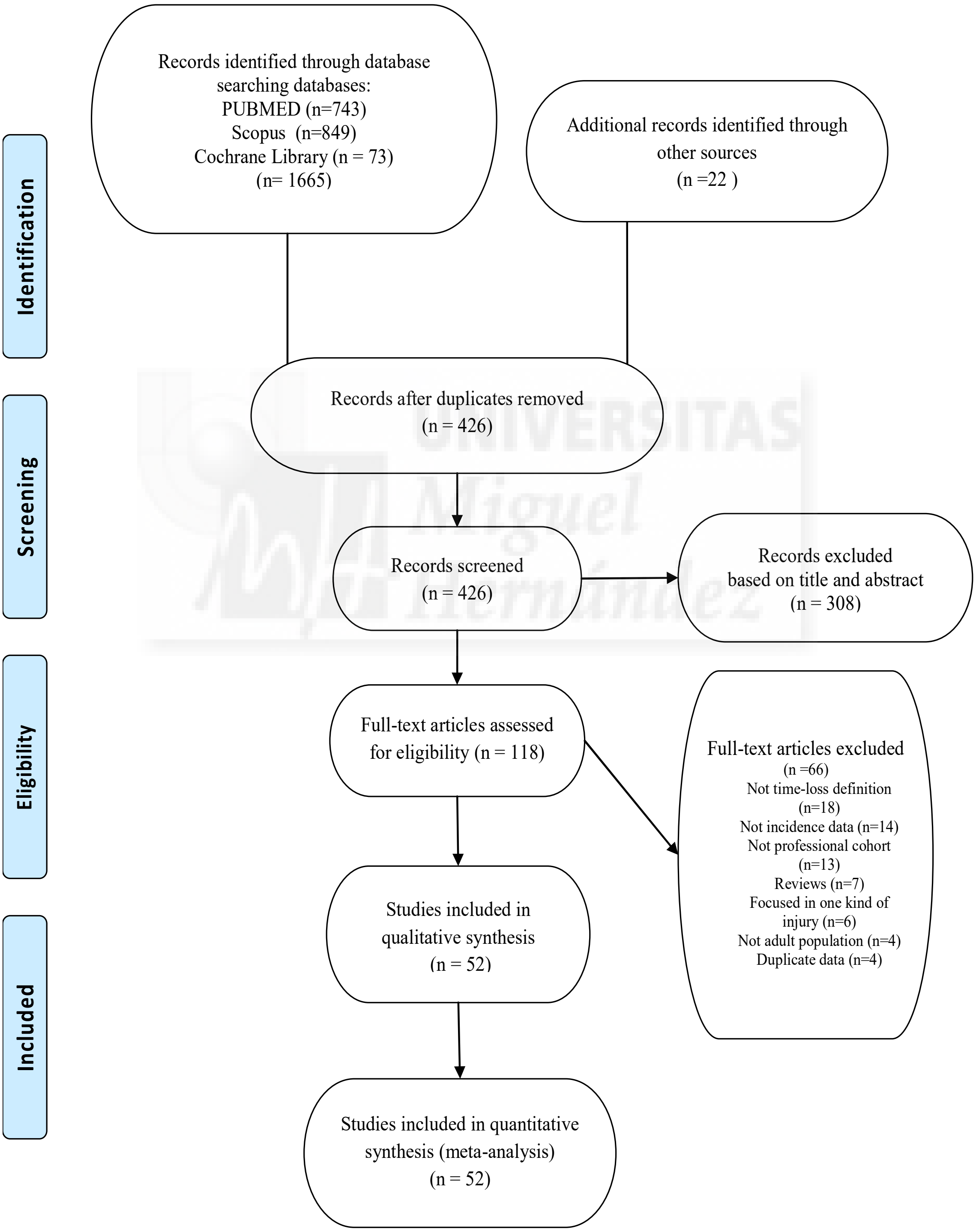


Figure 3.1. Flow chart of the selection of studies for the meta-analysis.

With regards to the methodological quality of the studies, the mean score obtained with the quality scale (range: 0–11) was 7.8 (minimum: 3, maximum: 11) (higher score indicates better quality). In general, more recent studies (published from 2007 to 2017) had a better methodological quality scores (8.4, 95% CI = 7.9 to 8.7) than older (published before 2007) studies (6.5, 95% CI = 6.2 to 7.1). Few of the studies selected satisfied items 4 (gives the inclusion and exclusion criteria),[22,29,38,41-43,50,71,77,88,179-181,184,185,187,189] five (describes injury history)[24,49,50,71,106,178,181,182,188,189] and 10 (indicates the number of participants with missing data and explain how this was addressed).[21,26,28,29,77,177,178,189] Contrarily, most studies (> 80%) described both the setting or participating locations (item 1)[16,18,20-24,26-29,33-43,49,50,71,77,83,88,99,100,106,174-194] and relevant dates (period of recruitment, exposure, follow-up, data collection) (item 2),[16,18,20-24,26,28,29,33,35-38,40-43,49,50,71,77,83,88,99,175-191,194] as well as verified injuries by an independent medical professional (item 8)[18,20-24,26-29,33-43,49,50,71,77,83,88,100,106,174,176-185,187-194] and/or classified injuries (severity, location and type of injury) (item 9).[16,18,20,22-24,26-29,33,34,36-43,49,50,71,77,83,88,99,106,174-185,187-194] The detailed data are presented in appendix 3.2.

Table 3.1. Characteristics of the studies included in the meta-analysis.

Reference	Study Duration*	Nº Teams	Exposure (hours)			Injuries			Incidence			STROBE quality
Country/Tournament		(Players)	Overall	Training	Match	Overall	Training	Match	Overall	Training	Match	
Almutawa et al., 2014 ^{F} *Yemen/Saudi NT – 2010*	4	1 (31)	1901.1	1785.6	115.5	18	8	10	9.5	4.5	86.6	High Quality
Almutawa et al., 2014 ^{F} *Qatar/Saudi NT – 2011*	4	1 (32)	1481.4	1382.4	99	26	16	10	17.6	11.6	101	High Quality
Andersen et al., 2004 Norway – 2000	39	- (330)	-	-	5742	-	-	121	-	-	21.5	High Quality
Arnason et al., 1996 Iceland – 1991	39	5 (84)	6854.84	-	-	85	-	-	12.4	-	-	High Quality
Arnason et al., 2004 Iceland – 1999	39	17 -	33839	-	-	244	-	-	7.2	-	-	High Quality
Arnason et al., 2005 Control Iceland – 2000	39	8 (144)	14617	11773	2844	96	22	74	6.6	1.9	26.9	Low Quality
Bjørneboe et al., 2010 AT Norway – 2004	156	14 -	74612	66563	8049	267	125	142	2.1	1.9	17.6	High Quality
Bjørneboe et al., 2010 Grass ^{F} Norway – 2004	156	14 -	186929	156002	30927	800	274	526	2.1	1.8	17	High Quality
Bjørneboe et al., 2014 ^{F} Norway – 2002	234	14 -	494157	-	-	2365	-	-	4.8	-	-	High Quality
Carling et al., 2010 ^{F} France – 2005	156	1 (124)	-	-	3246.7	-	-	130	-	-	40.5	High Quality
Carling et al., 2015 ^{F} France – 2009	195	1 (130)	52513	47000	5513	372	141	231	7.1	3	41.9	High Quality
Dupont et al., 2010 ^{F} Scotland – 2007	78	1 (32)	18495	16339	2156	165	60	105	8.9	3.7	48.7	High Quality
Dvorak et al., 2007 ^{F} *Germany/WC - 2006*	4	32 (736)	-	-	2112	-	-	145	-	-	45.9	Low Quality
Dvorak, et al., 2011 ^{F} *South Africa/WC – 2010*	4	32 (736)	15206	13160	2046	140	58	82	9.2	4.4	40.1	High Quality
Eirale et al., 2012 ^{F} *Qatar NT – 2007*	68	1 (36)	10043	9482	561	78	41	37	7.8	4.3	65.4	High Quality
Eirale et al., 2013a ^{F} Qatar – 2008	39	10 (230)	36166	32844	5793	217	133	84	6	4.4	14.5	High Quality
Eirale et al., 2013b ^{F} Qatar – 2008	117	5 (527)	164434	142230	22206	826	462	364	5	3.1	16.4	High Quality

Ekstrand & Tropp, 1990 (a) Sweden	39	9 (135)	30554	23241	7313	261	107	159	8.6	4.6	21.8	Low Quality
Ekstrand & Tropp, 1990 (b) Sweden	39	12 (180)	33652	24499	7219	288	125	135	8.5	5.1	21.8	Low Quality
Ekstrand et al., 2004 *Sweden NT – 1991*	234	1 -	7245	6235	1010	71	40	31	10	6.5	30.3	Low Quality
Ekstrand et al., 2010 ^{F} European teams – 2001	273	23 (1065)	566000	475000	91000	4483	2546	1937	8	4.1	27.5	High Quality
Ekstrand et al., 2011a ^{F} European teams – 2003	195	15 -	218265	-	-	1791	-	-	8.2	-	-	High Quality
Ekstrand et al., 2011b ^{F} European teams – 2001	613	51 (2299)	1175000	-	-	9275	-	-	7.9	-	-	High Quality
Ekstrand et al., 2013 ^{F} European teams – 2001	429	160 (1743)	1057201	888249	168952	8029	3483	4546	7.6	4	26.7	High Quality
Engström et al., 1990 Sweden	39	3 (64)	17000	14592	3136	85	44	41	5	3	13	Low Quality
Fünten et al., 2014 ^{F} Germany – 2008	34	7 (254)	48324	42932	5483	300	141	159	6.2	3.3	29	High Quality
Gouttebarge et al., 2016 ^{F} Australia – 2008	195	49 (1127)	-	-	10791	-	-	845	-	-	78.3	High Quality
Gustafsson, 2011 (a) ^{F} European teams – 2001	312	113 (1349)	382219.1	310460.5	71758.6	2725	964	1761	7.1	3.1	24.5	High Quality
Gustafsson, 2011 (b) ^{F} European teams – 2001	312	113 (1644)	377937.9	328835.7	49102.2	2662	1387	1275	7	4.2	26	High Quality
Hägglund et al., 2003 (a) Sweden – 1982	39	8 (118)	27966	21476	6490	236	99	137	8.3	4.6	20.6	High Quality
Hägglund et al., 2003 (b) Sweden – 2001	39	14 (310)	93310	81840	11470	715	421	294	7.8	5.2	25.9	High Quality
Hägglund et al., 2005 Denmark – 2001	39	8 (188)	27321	23095	4226	395	271	124	14.4	11.8	28.2	High Quality
Hägglund et al., 2006 (a) ^{F} Sweden – 2001	39	12 (263)	78597	68849	9748	601	349	252	7.6	5.1	25.9	High Quality
Hägglund et al., 2006 (b) ^{F} Sweden – 2002	39	12 (262)	77270	66973	10397	588	352	236	7.6	5.3	22.7	High Quality
Hägglund et al., 2009a ^{F} Sweden- 2005	39	11 (239)	71361	62315	9046	548	294	254	7.7	4.7	28.1	High Quality
Hägglund et al., 2009b (a) ^{F} Austria-Switzerland/EC - 2008	2	16 (367)	5368	4310	1058	56	12	44	10.4	2.8	41.6	High Quality

Hägglund et al., 2009b (b) F Portugal/EC U21 – 2006	1	8 (176)	1589	1076	513	22	5	17	13.8	4.6	33.1	High Quality
Hägglund et al., 2009b (c) F Netherland /EC U21 - 2007	1	8 (182)	2321	1774	548	25	6	19	10.8	3.4	37.4	High Quality
Hägglund et al., 2009b (d) F Polonia/EC U19 - 2006	1	8 (144)	1253	762	490	8	0	8	6.4	0	16.3	High Quality
Hägglund et al., 2009b (e) F Austria/EC U19 - 2007	1	8 (147)	1158	654	504	15	1	14	13	1.5	27.8	High Quality
Hägglund et al., 2009b (f) F Czech Republic/EC U19 - 2008	1	8 (145)	1461	957	504	15	2	13	10.3	2.1	25.8	High Quality
Hägglund et al., 2016 (a) F European teams - 2001	546	43 (6956)	1613792	1356420	257372	11581	5154	6512	7.2	3.8	25.3	High Quality
Hägglund et al., 2016 (b) F European teams - 2001	351	19 (2014)	521626	453150	68476	3836	2220	1609	7.4	4.9	23.5	High Quality
Hassabi et al., 2010 Iran - 2005	16	1 (21)	2610	-	-	50	-	-	19.2	-	-	Low Quality
Hawkins & Fuller, 1996 United States/WC - 1994	4	24 (412)	-	-	1657	-	-	104	-	-	67.8	Low Quality
Hawkins & Fuller, 1999 England - 1994	407	4 (138)	68000	55000	15096.5	578	187	391	8.5	3.4	25.9	High Quality
Junge & Dvořák, 2013 FIFA U20 WC	21	- -	-	-	12012	-	-	314	-	-	26.1	High Quality
Junge & Dvořák, 2013 FIFA Confederation Cups	10	- -	-	-	2640	-	-	51	-	-	19.3	High Quality
Junge & Dvořák, 2013 FIFA World Club Cups	13.5	- -	-	-	2541	-	-	81	-	-	31.9	High Quality
Junge & Dvořák, 2015 F Brazil/WC - 2014	4	32 (736)	-	-	2046	-	-	60	-	-	29.3	High Quality
Junge et al., 2004a (a) France/WC - 1998	4	32 (736)	-	-	2046	-	-	149	-	-	72.8	High Quality
Junge et al., 2004a (b) Sydney/OG - 2000	2	16 -	-	-	1023	-	-	52	-	-	50.8	High Quality
Junge et al., 2004b Korea-Japan/WC - 2002	4	16 (736)	-	-	2112	-	-	107	-	-	50.7	Low Quality
Kristenson et al., 2013 F Norway-Sweden -2010	78	33 (1044)	367490	318568	48922	2241	1178	1063	6.1	3.7	21.7	High Quality
Lee et al., 2014 F Hong-Kong - 2010	39	7 (152)	39768.5	37143	2717	296	130	166	7.4	3.5	61.1	High Quality

Mallo et al., 2011 [F] Spain - 2003	156	1 (129)	28694	24509	4185	313	129	184	10.9	5.2	44.1	High Quality
Nielsen & Yde, 1989 Denmark - 1896	39	2 (34)	31020	28695,7	2324,3	37	-	-	3.5	2.3	18.5	Low Quality
Noya et al., 2014a [F] Spain - 2008	39	11 (301)	161602.7	153567.2	8035.5	891	579	312	5.5	3.8	38.8	High Quality
Noya et al., 2014b [F] Spain - 2008	39	16 (427)	228743	216705	12038	1293	769	524	5.7	3.6	43.5	High Quality
Pedrinelli et al., 2013 Argentina/AmC - 2011	4	12 (276)	-	-	891	-	-	63	-	-	70.7	High Quality
Poulsen et al., 1991 Denmark - 1986	39	1 (19)	4199	3440	759	29	14	15	6.9	4.1	19.8	Low Quality
Reis et al., 2015 [F] Brazil	48	1 (48)	16077	15120	957	70	29	41	4.4	2.4	42.8	High Quality
Shalaj et al., 2016 [F] Kosovo - 2003	39	11 (143)	36833	31998	4834	272	101	171	7.4	3.2	35.4	High Quality
Stubbe et al., 2015 [F] Netherland - 2009	39	8 (217)	46194	41012	5182	286	116	170	6.2	2.8	32.8	High Quality
Waldén et al., 2005 European teams - 2001	47	11 (266)	69707	58149	11558	658	298	360	9.4	5.8	30.5	High Quality
Waldén et al., 2007 (a) [F] Portugal/EC - 2004	3	16 (368)	4742	3694	1048	45	7	38	10.1	2.1	36	High Quality
Waldén et al., 2007 (b) [F] Ireland/ EC U19 - 2005	5	8 (144)	1394	899	405	17	2	15	13.4	2.9	30.4	High Quality
Waldén et al., 2013 (a) [F] European teams - 2001	351	20 (1100)	595498	499238	96260	4699	2035	2664	7.9	4.1	27.7	High Quality
Waldén et al., 2013 (b) [F] European teams - 2001	351	5 (257)	178066	149948	28118	1250	528	722	7	3.5	25.7	High Quality

[F] Study was implemented according to the 2006 consensus statement for epidemiological studies in soccer.

(a);(b);(c): indicate different cohorts in the same study.

a; b: indicate different studies in the same year.

*: study duration expressed in number of weeks.

AT: artificial turf; WC: World cup; EC: European cup; AmC: American cup; OG: Olympic games; NT: national team.

3.4.2. Findings meta-analysis

In the different meta-analyses carried out, the effect sizes exhibited a moderate to large heterogeneity (based on the Q statistics and the I^2 indices), supporting the decision of applying random-effects models.

Neither weeks of follow-up, year of publication of the study, age, STROBE score and number of teams' variables had an impact on injury incidence rates and hence, the subsequent sub-analyses were not adjusted to these variables.

3.4.3. Injury incidence: overall, training and match

Forty-two studies (56 cohorts) reported overall injury incidence,[16,18,20-24,29,33-39,41-43,49,50,71,77,83,88,99,100,174,176-189,194] 36 studies (50 cohorts) reported training injury incidence,[16,18,21-24,29,33,35-39,42,49,50,71,77,88,99,100,174,176-189] and 46 studies (63 cohorts) reported match injury incidence[16,18,21-24,26-29,33,35-40,42,49,50,71,77,88,99,100,106,174-193] that could be combined in the meta-analysis. These studies comprised 67335 (overall), 17815 (training) and 21844 (match) injuries, among more than 28077, 24714 and 29191 professional male soccer players, respectively.

The random effect models for injury incidence showed an overall incidence of 7.7 injuries per 1000 hours of exposure (95% CI = 7.1 to 8.4), a training incidence of 3.9 injuries per 1000 hours of training exposure (95% CI = 3.4 to 4.3) and a match incidence of 32.9 injuries per 1000 hours of match exposure (95% CI = 29.2 to 36.5). Figures 3.2-3.4 display a summary of the reported overall, training and match injury incidence rates of the analysed studies, respectively. The injury incidence rate during matches was most likely higher (100% likelihood) than the training injury incidence rate.

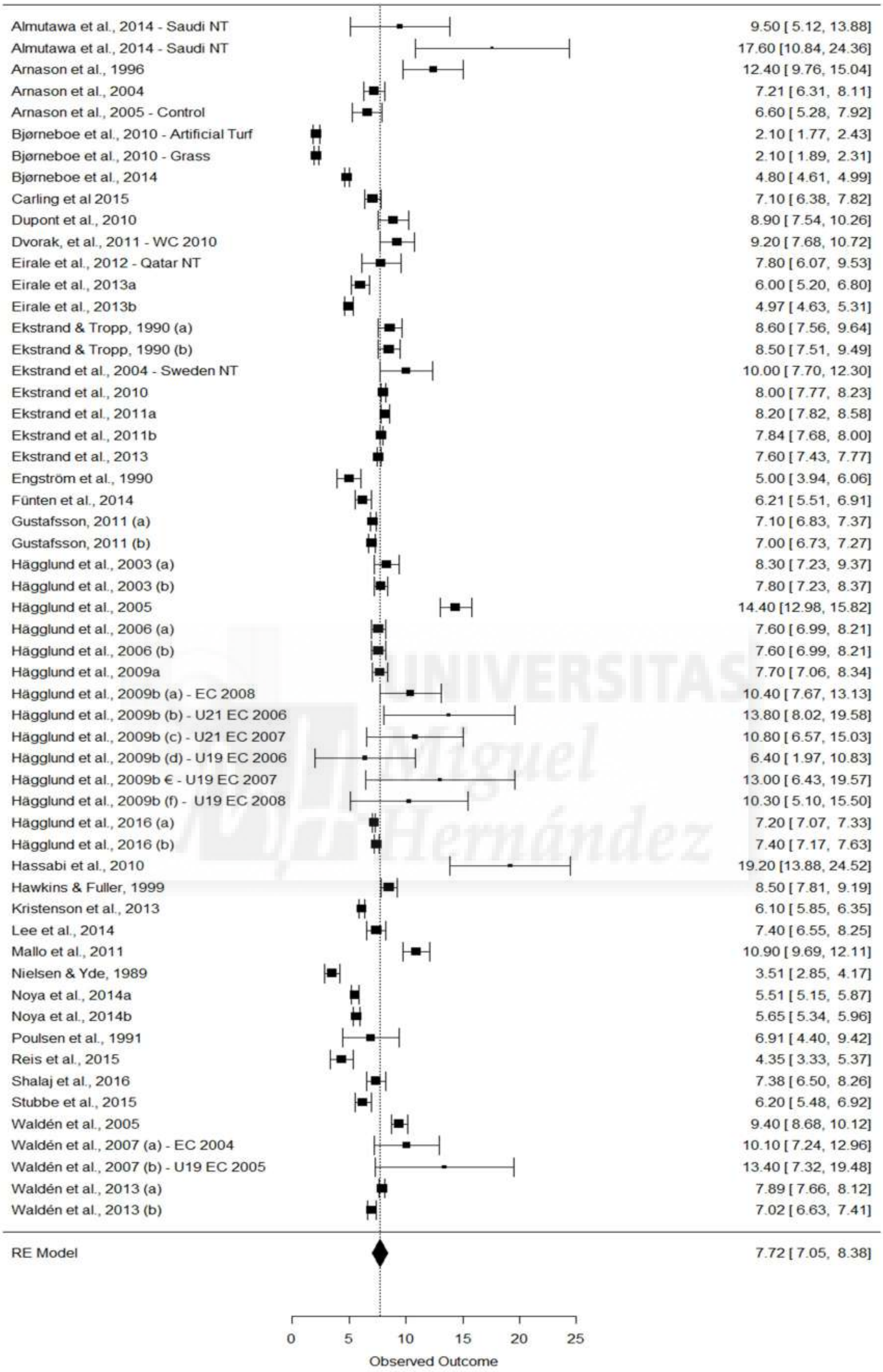


Figure 3.2. Overall injury incidence with 95% confidence intervals. EC: European cup; NT: national team; WC: World cup.

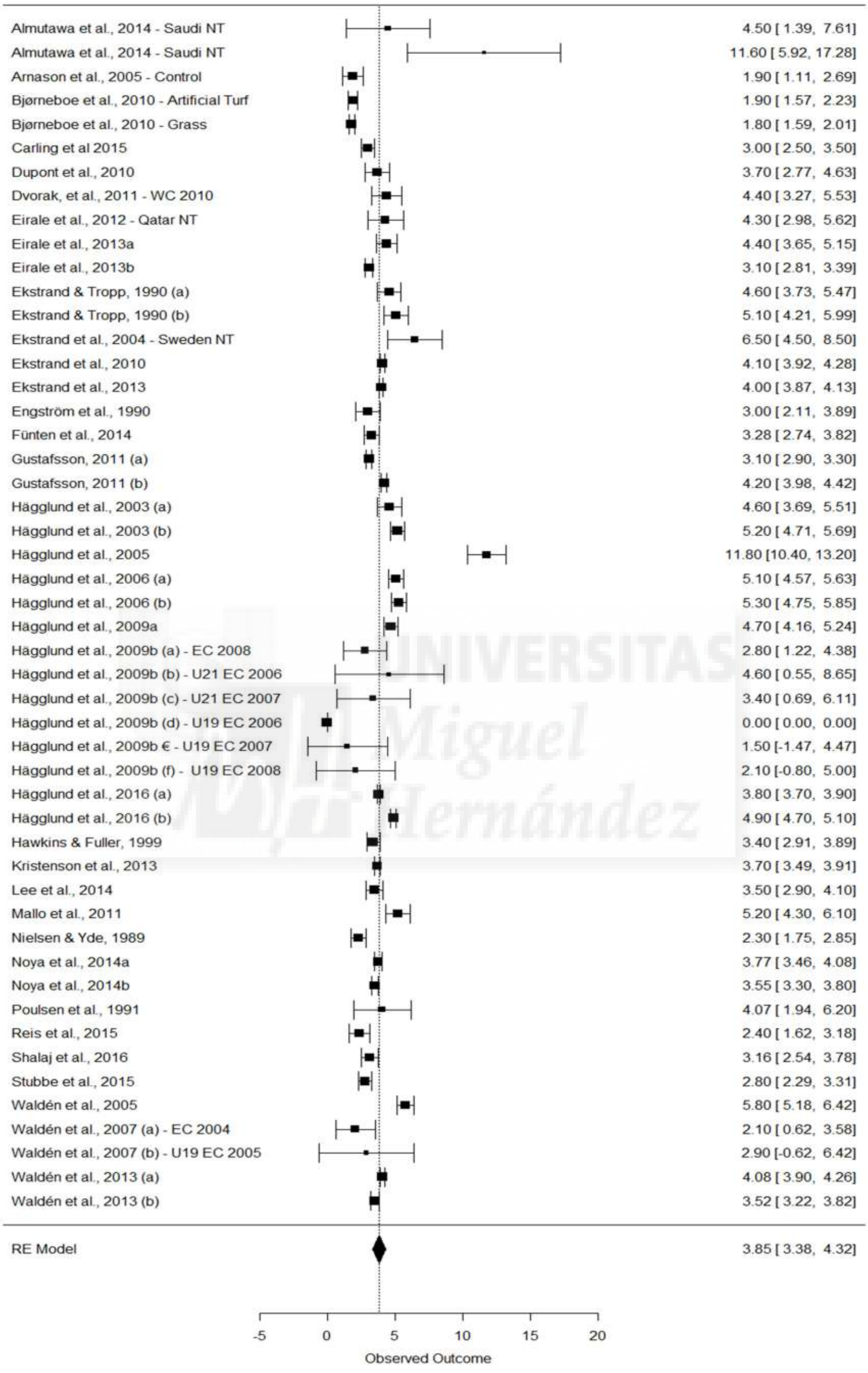


Figure 3.3. Training injury incidence with 95% confidence intervals. EC: European cup; NT: national team; WC: World cup.

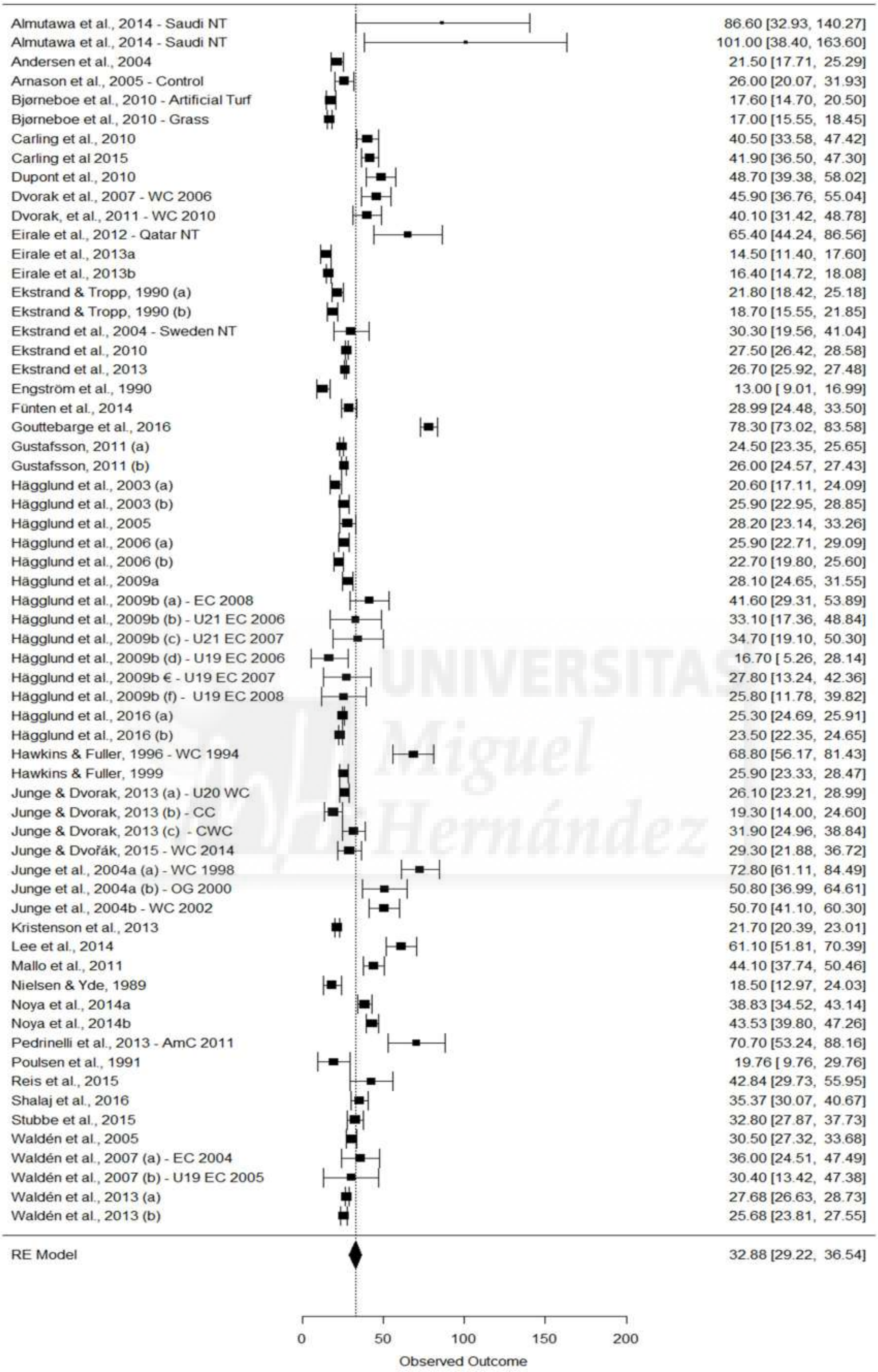


Figure 3.4. Match injury incidence with 95% confidence intervals. AmC: American cup; CC: Confederations cup; CWC: Clubs World cup; EC: European cup; NT: national team; OG: Olympic Games; WC: World cup.

3.4.4. Location of injury

Thirty studies (34 cohorts) reporting the main groupings of injury location and lower extremities region categories according to Fuller et al.,[19] were pooled in the meta-analysis.[16,18,20,22-24,29,33,34,36,38,39,43,49,71,77,83,88,174,177-179,181,183-186,188,189,194] Lower extremity injuries had the highest incidence rates (6.5 per 1000 hours of exposure, 95% CI = 5.7 to 7.3) compared to the other body regions. The trunk was the second most commonly injured region (0.4 per 1000 hours of exposure, 95% CI = 0.3 to 0.5), upper extremity was the third most commonly injured region (0.2 per 1000 hours of exposure, 95% CI = 0.1 to 0.3) followed by head and neck injuries (0.2 per 1000 hours of exposure, 95% CI = 0.1 to 0.3) and undefined/other injuries had the lowest incidence rates (0.04 per 1000 hours of exposure, 95% CI = 0.01 to 0.07). The incidence rate of lower extremity injuries was most likely higher (100% likelihood) than other location, whereas trunk injuries occurred more often (100% likelihood) than upper extremity and head and neck injuries. Differences between upper extremity injury incidence rate and head and neck injury incidence rate were trivial. There were no significant differences between the remaining paired combinations.

Regarding lower extremity injuries, six anatomical regions were analysed. The mean incidence per 1000 player hours of exposure with 95% CIs were in descending order: thigh (1.8, 1.6 to 2.1), knee (1.2, 1.1 to 1.4), ankle (1.0, 0.9 to 1.1), hip/groin (0.9, 0.7 to 1.0), lower leg/Achilles tendon (0.8, 0.7 to 0.9), and foot/toe (0.4, 0.3 to 0.5). In term of paired-comparisons, thigh injuries occurred significantly more frequently (100% likelihood) than injuries in other lower extremity regions. Foot incidence rate was most likely lower (100% likelihood) than other lower extremity regions. Knee injury rates were most likely higher (100% likelihood) than lower leg rates, as well as very likely higher (98% likelihood) and likely higher (95% likelihood) than hip and ankle incidence rates, respectively. Otherwise, ankle injury rate were likely higher (85% likelihood) than lower legs/Achilles tendon. There were no significant differences between the remaining paired combinations (Figure 3.5).

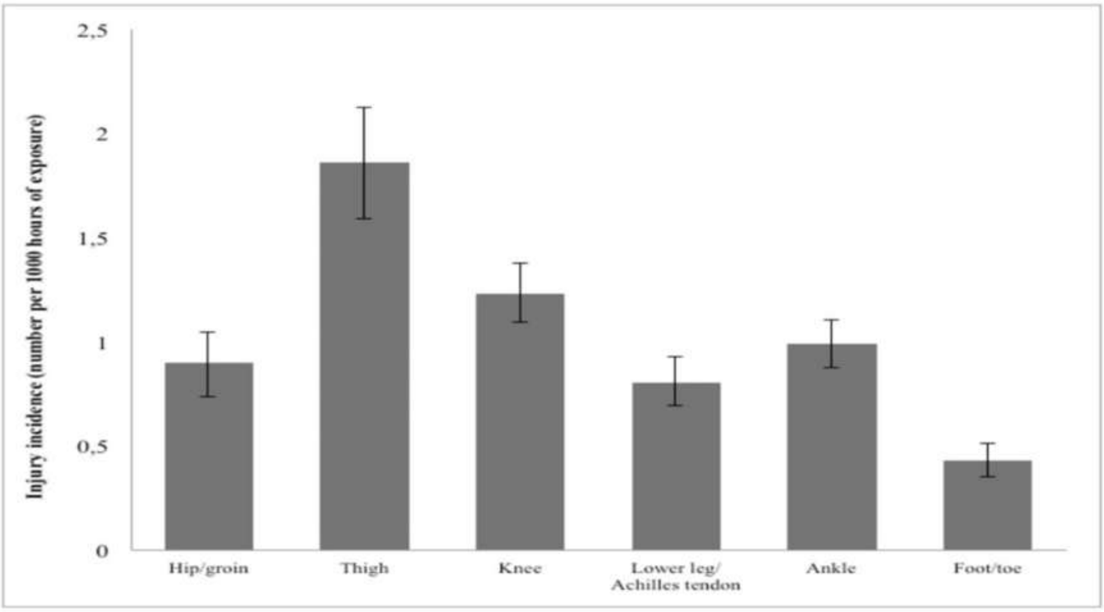


Figure 3.5. Injury incidence rates (with 95% confidence intervals) by location of lower extremity injuries.

3.4.5. Type of injury

An analysis was undertaken to determine the most frequent type of injury sustained in professional male soccer players. Twenty-nine studies (33 cohorts) were included in the pooled analysis.[16,18,20,22-24,29,33,34,36,38,39,43,49,83,88,174,177-179,181-186,188,189,194] The mean incidence is presented per 1000 hours of exposure with 95% CIs. The most common type of injury was muscle/tendon (4.4, 3.9 to 4.9), followed by contusions (1.3, 1.1 to 1.5), undefined/other injuries (0.8, 0.5 to 1.1), joint (non-bone) and ligament (0.3, 0.2 to 0.4), fracture and bone stress (0.2, 0.1 to 0.3), central/peripheral nervous system (0.05, 0.03 to 0.07) and the least common injury type was laceration and skin lesions (0.04, 0.02 to 0.06). Muscle/tendon injury incidence rates were most likely higher than other types of injuries rates (100% likelihood). Furthermore, skin lesions were most likely lower (100% likelihood) than all types of injuries with the exception of central/peripheral nervous system injuries. Likewise, contusions incidence rate was most likely higher (100% likelihood) than other types of injuries, with the exception of the muscle/tendon injuries rate. Finally, both joint (non-bone) and ligament as well as fracture and bone stress injuries incidence rates were most likely higher (100% likelihood) than central/peripheral nervous system injuries rate (Figure 3.6).

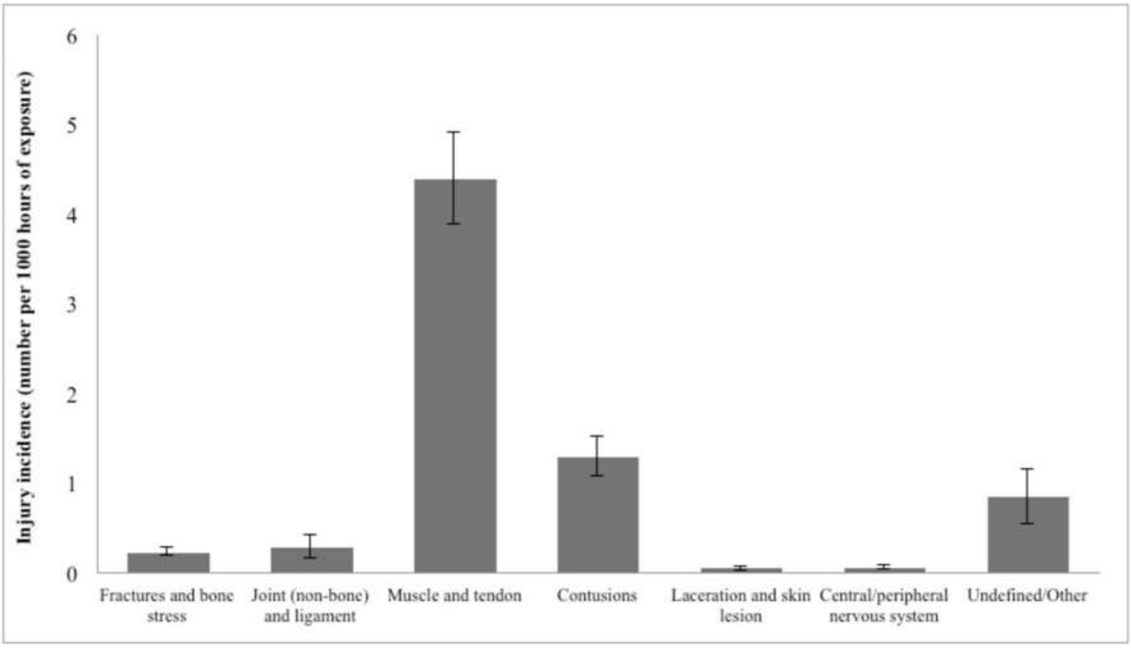


Figure 3.6. Injury incidence (with 95% confidence intervals) by type of injury.

3.4.6. Severity of injury

Concerning severity of injuries, 26 studies (37 cohorts) were included in the pooled analysis.[16,18,21-24,29,36-38,42,49,71,77,88,174,176-179,181-185,188,189] Minimal injuries (2.8 per 1000 hours of exposure, 95% CI = 2.3 to 3.4) were the most usual injuries, followed by moderate (2.1 per 1000 hours of exposure, 95% CI = 1.8 to 2.4), minor (1.8 per 1000 hours of exposure, 95% CI = 1.5 to 2.0) and severe (0.8 per 1000 hours of exposure, 95% CI = 0.6 to 0.9) injuries.

Comparisons between each severity level showed that the minimal injuries rate was most likely higher (100% likelihood) than other severities. Minor injuries rate was most likely higher (100% likelihood) than severe injuries and possibly lower (73% likelihood) than moderate injuries rates. Moderate injuries incidence rate was most likely higher than severe injuries incidence rates (100% likelihood) (Figure 3.7).

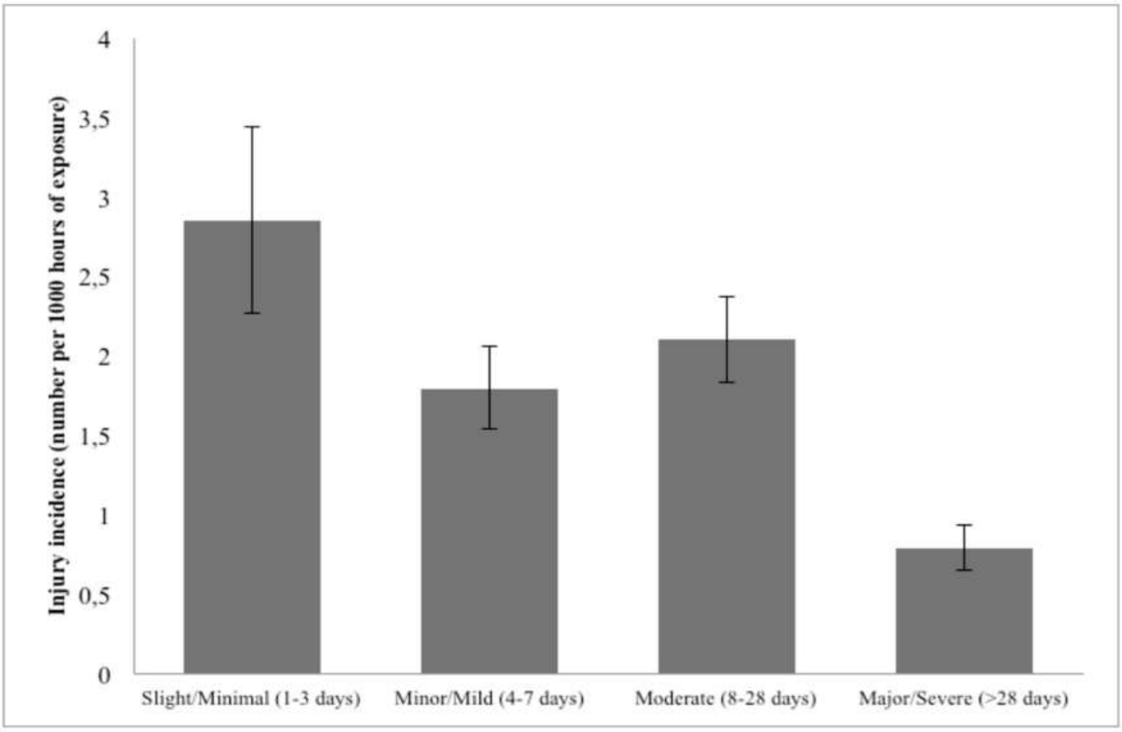


Figure 3.7. Injury incidence (with 95% confidence intervals) by severity of injury.

Nineteen studies (20 cohorts) reported an average injury time-loss of 13.8 days per injury and player.[16,18,21,22,24,34,36,38,41,49,176,179-181,183,185,187,189,191]

3.4.7. *Overuse vs. traumatic injuries*

Twenty-four studies (32 cohorts) were involved in the meta-analysis in order to compare overuse injuries versus traumatic injuries.[16,20,22,23,29,36-38,43,49,71,77,83,88,178,179,181,183-186,188,189,194] The incidence in traumatic injuries (5.3, 4.4 to 6.3) was most likely higher (100% likelihood) than in overuse injuries (2.4, 1.9 to 2.9).

3.4.8. *New vs. recurrent injuries*

Twenty studies (26 cohorts) were included in an analysis which comparing the incidence of new versus recurrent injuries.[16,18,22-24,35-38,49,71,88,174,177-179,183,184,186,189,194] The incidence rate of new injuries (6.7 per 1000 hours of exposure, 95% CI = 6.0 to 7.5) was most likely higher (100% likelihood) than recurrent injuries incidence rate (1.2 per 1000 hours of exposure, 95% CI = 0.9 to 1.6).

3.4.9. Level of play

For level of play, 52 studies were divided into two groups: a) national level - studies in professional soccer clubs; and b) international level - studies in national teams. Thirty-six,[18,20-24,33-36,38,39,41,43,49,50,71,77,83,88,99,100,174,176,178,180-189,194] 32[18,21-24,29,33,35,36,38,39,49,50,71,88,99,100,174,176,178,180-189,194] and 35 studies[18,21-24,29,33,35,36,38-40,49,50,71,88,99,100,174,176,178,180-191,194] carried out in soccer clubs reported overall, training and match incidence rates, respectively. On the other hand, six,[16,29,37,42,177,179] six[16,29,37,42,177,179] and 13[16,26-29,37,42,106,175,177,179,192,193] studies in national teams reported overall, training and match incidence rates, respectively. Incidence rate in international level was most likely higher (100% likelihood) than national level (9.8, 95% CI = 8.7 to 10.8 vs. 7.2, 95% CI = 6.5 to 7.9, respectively). In particular, the mean incidence rates in training and match were in descending order: international match: (41.1, 95% CI = 33.9 to 48.2), national match (29.2, 95% CI = 25.3 to 33.0), national training (4, 95% CI = 3.5 to 4.5) and international training (3.4, 95% CI = 2.2 to 4.7). The incidence rate during international matches was likely higher (99% likelihood) than during national matches. There were no statistical relevant differences in incidence rates between international and national training injuries.

3.4.10. Trends in injury risk over time

Trend in incidence risk reported by epidemiologic studies before and after the publication of the “consensus statement on injury definitions and data collection procedures”[19] were compared. Studies published after this consensus showed slightly lower overall (7.5 injuries per 1000 hours of exposure [95% CI = 6.7 to 8.2] vs. 8.3 injuries per 1000 hours of exposure [95% CI = 6.9 to 9.7]) and training (3.5 training injuries per 1000 hours of exposure [95% CI = 3.2 to 3.9] vs. 4.8 injuries per 1000 hours of exposure [95% CI = 3.4 to 6.3]) injuries. However, match injury incidence rate was higher in studies after the consensus (33.5 injuries per 1000 hours of exposure [95% CI = 29.3 to 37.6] vs. 31.3 injuries per 1000 hours of exposure [95% CI = 23.3 to 39.3]. Likewise, we compared injury incidence rates among three decades (1989-1999; 2000-2009; 2010-2016) and the incidence of match injuries showed an increasing trend, while a decreasing trend in training injuries were observed (Table 3.2).

Table 3.2. Trend in injury incidence (expressed in 1000 hours of exposure) over the last three decades.

Period of time	Mean	95% CI	
		Lower	**Upper**
1989-1999 (8 cohorts)			
▪ Overall incidence	7.5	5.4	9.6
▪ Training incidence	4.1	1.9	6.2
▪ Match incidence	26	12.9	39.1
2000-2009 (23 cohorts)			
▪ Overall incidence	9.2	8.0	10.2
▪ Training incidence	4.2	2.9	5.5
▪ Match incidence	32.4	27.2	37.6
2010-2017 (38 cohorts)			
▪ Overall incidence	7	6.2	7.8
▪ Training incidence	3.6	3.3	3.9
▪ Match incidence	34.7	29.5	40.0

CI: confidence intervals.

3.5. Discussion

The main purpose of the current study was to carry out a systematic review and meta-analysis of injury incidence in professional male soccer as reported in the literature, as well as to make magnitude-based inferences regarding location of injuries, type of injuries, severity of injuries, overuse vs. traumatic injuries, new vs. recurrent injuries, level of play and trend in injury incidence over time. To address this purpose, data from 52 epidemiologic articles (49 prospective and three retrospective studies) which used comparable methodologies were pooled and meta-analysed, producing estimates of injuries that more accurately reflect the injury incidence present amongst professional/elite soccer players than data provided in individual studies.

This novel meta-analysis of injury incidence and severity in professional male soccer players underlines the moderate to high overall injury incidence rate of this sport modality (7.7 injuries per 1000 hours of exposure), and its similitude with other injury incidence rates provided in individual studies for amateur soccer players (9.6 injuries per 1000 hours of exposure)[195] and other professional team sports such as rugby (7.9 injuries per 1000 hours of exposure).[196] Fortunately, although injuries occur frequently in professional soccer players, the majority appear to be of minimal severity (2.8 per 1000 hours of exposure, 95% CI = 2.3 to 3.4). These injury incidence rates resulted on average in 13.8 days lost from soccer activities (trainings and matches).

In line with what happens in most team sports (e.g. basketball,[31] handball,[32] rugby[30]), match/game injury incidence in soccer (32.9 injuries per 1000 hours of exposure [95% CI = 29.2 to 36.5]) was notably higher (eight times) than the injury rate obtained for training sessions (3.9 injuries per 1000 hours of exposure [95% CI = 3.4 to 4.3]). A number of studies have attributed these differences in injury incidence rates between match and training to several factors, including: the higher physical demands of players during matches in comparison with training sessions, the variability and uncertainly generated in the players for competing against rivals, the number of contacts and collisions accounted during the matches, and the fatigue generated during the course of the match.[18,41,197,198] Although yet under debate, it has been suggested that training session design (i.e. work-load, intensity, duration), when possible, should mimic match demands so that players would be better ready in terms of robustness for what they will face during matches[199]. Likewise, the congested competitive calendar prevailing in professional soccer (specially in top teams) might result in players developing chronic sub-optimal readiness situations (i.e. overtraining) caused among others, by an insufficient post-competition recovery and/or accumulated fatigue. It potentially could increase the risk of injury during matches, especially during the last weeks of the season when domestic and European competitions approach their climax and 2-3 matches per week are common.[68,71,151] In relation with this latter circumstance, increasing the squad size (e.g. 25 players) might be an interesting measure to counteract the negative effects of the match congestion on injury risk, as coaches could rotate players, providing a greater time-interval between consecutive matches that potentially would allow for better post-competition recovery.

Although the limited number of studies published did not permit us to describe the pattern of the injury incidence during the course of a soccer match, the current evidence shows that injuries tend to occur more frequently toward the end of each half.[18,25,28,41,76,106,107] The finding of a higher incidence of injuries in the second part of each half in comparison with other match periods may indicate that fatigue is implicated in injury aetiology, however, factors contributing to this (e.g. hydration, nutrition, neuromuscular compromise and biomechanical alterations to technique) require further investigation. Perhaps, allowing more substitutes on the bench at matches might help to reduce the number of injuries accounted toward the end of each half. Similarly, increasing the number of resting intervals during matches (from the current 15 min-resting interval to two or even four [as happen in basketball] short-rests intervals) or allowing coaches to request time-outs might also help to reduce the impact of fatigue on injury risk.

With respect to the location of soccer-related injuries, as expected, lower extremity injuries were, by far, the most frequent, with an incidence of 6.5 injuries per 1000 hours of exposure (95% CI = 5.7 to 7.3). The thigh was the anatomical region of the lower extremity where injuries occurred significantly more frequently, with an incident rate of 1.8 injuries per 1000 hours of exposure (95% CI = 1.6 to 2.1), followed by the knee, with an incidence rate of 1.2 injuries per 1000 hours of exposure (95% CI = 1.1 to 1.4). Furthermore, and also as expected, the most common type of injury was muscle/tendon injuries (4.4 injuries per 1000 hours of exposure [95% CI = 3.9 to 4.9]). Due to the lack of studies reporting incidence rates separately for different muscle groups (e.g. gluteus, hamstrings, quadriceps, abductors, adductors, triceps surae), a sub-analysis aimed at identifying the most injured muscle group was not possible to carry out. However, previous epidemiological studies have consistently reported that hamstring muscles are the most frequent injured in professional soccer players.[41,103,189,197,200] It has been estimated that a soccer team with a 25 player-squad typically suffers about 5–6 hamstring injuries each season equivalent to more than 80 days involving football activities (trainings or matches) lost due to injury.[18] This quantity of time-loss could be significant because players' side-lined due to injury limit the possibility of optimal performance by the team. Therefore, coaching, medical and fitness staff should work together as a team in order to design strategies aimed at reducing the number of hamstring injuries. According to Ekstrand, Waldén, & Hägglund,[48] these strategies should not only address the proposed internal hamstring injury risk factors (e.g. eccentric strength deficits,[118,123,125,201] poor flexibility,[53,56,202-204] altered muscle architecture[205-207]) but also external factors (e.g. player load and match frequency[20,208,209] or stability of the club in terms of coaching, medical staff and management).

The findings of this study also reported that most of the soccer-related injuries had a traumatic mechanism (injuries with sudden onset and known cause), with an incidence rate of 5.3 injuries per 1000 hours of exposure, twofold that the incidence reported by overuse injuries (injury with insidious onset and no known trauma), 2.4 injuries per 1000 hours of exposure. In particular, tackled or colliding with a rival (e.g. during a jump) appear to be the most common injury incidents, representing approximately 50% of all the traumatic injuries, followed closely by the injury incidents caused during non-contact actions such as sprinting and cutting, reaching the 30% of all cases of traumatic injuries.[44,62] As it has been documented for young players,[210-213] the application of soccer-specific neuromuscular training programmes with the aim of optimizing players′ motor competency, joint stability (e.g. knee, ankle and core) and delaying the onset of fatigue might reduce the relative risk of injury due to acute overload of soft tissues (ligaments, tendons and muscles). Similarly, reducing the number of contact injuries might be achieved by changing the rules of the sport, for example punishing with more severity tackles and voluntary collisions, as well as by delivering formative programs to coaches and players aimed at warning about the risk that these actions have. It might help to reduce the incidence of the contact injuries. Likewise, teaching players how they should face tackles from a technical perspective might also be an interesting measure to reduce contact injuries. The efficacy of the just-mentioned measures for reducing the number of contact injuries in professional soccer is still unknown and requires further study.

Although the recurrent injury incidence rate is lower than the new injuries rate (1.2 vs. 6.7 injuries per 1000 hours of exposure) and that a slight decreased has been recently shown,[35] its magnitude still being significant. Therefore, the scientific community should continue making efforts in order to improve decision-making process for a safe return to play. In particular, future studies should extend our current knowledge further by employing and designing learning algorithms or artificial intelligence-based models that allow the identification of when a player is fully and effectively rehabilitated before returning to play.

The results of this study also highlight that the incidence rate during international matches (41.1 injuries per 1000 hours of exposure) was higher than during national matches (29.2 injuries per 1000 hours of exposure). The higher density of matches played, the mental stress and anxiety generated in the players and the fact that international competitions are usually played during summer periods (at the end of a long season where accumulated fatigue may play a part and during hot and dry climate conditions) have been suggested as contributing factors for this increase in the number of injuries.[16,176,182,191]

Finally, the present meta-analysis has shown the incidence of match injury, unlike training injury incidence, has increased after the publication of the consensus statement in 2006. Therefore more robust and consistent methods for determining injury incidence probably will provide us with a more accurate reflection of injury incidence. Likewise, time-trend analysis shows a slight upward tendency (non-statistically significant) in match injury incidence over the last three decades. However, data available from eligible studies did not permit the identification of whether the increase in the number of match injuries is followed by an increase in their severity. Consequently, future studies should be carried out in order to clarify the issue.

3.5.1. Limitations

Similar to other meta-analysis conducted in sport medicine settings,[30,167,214] methodological limitations were associated with many of the older studies included in this review, namely variations in injury and severity definitions, and a lack of uniform data collection methods. Since the 2006 consensus statement,[19] the methodological quality of published studies has improved, allowing for more effective interpretation and comparison of findings across studies. Nonetheless, it is difficult to ensure consistency in reporting and data collection practices across studies and teams. A recognised limitation of the present study is that the sample size of studies included was not sufficient to investigate interactive effects within factors (e.g. playing position by level of play) or whether injury rates are associated with a violation of the match rules (a variable that has not been thoroughly explored).

3.6. Conclusions

Professional male soccer players are exposed to a substantial risk of sustaining injuries, especially during matches. Although most injuries had a traumatic mechanism (injuries with sudden onset and known cause), fortunately most of them appear to be of minimal severity. As might be expected the lower extremity is more frequently injured, and the most common type of injury is muscle/tendon strains. Recurrent injuries were less frequent than new injuries, although re-injury rates have implications for return to play management. There appears to be a slight upward tendency (non-statistically significant) in match injury incidence over the last three decades. Future studies should focus on introducing and evaluating preventative measures that target the most common diagnoses, namely, muscle/tendon injuries highlighted in this meta-analysis, in order to reduce the injury burden within male professional soccer players.

3.7. Appendixes

Appendix 3.1. Moderator variables codded.

General study descriptors

- Authors
- Year of the study
- Country/Tournament
- Sampling time (number of seasons)

Description of the study population

- Sample size
- Number of teams
- Age
- Level of play (club or national team)

Epidemiological descriptors

- Injury definition
- Number of injuries (total, match and training)
- Exposure time (total, match and training)
- Incidence (total, match and training)
- Injury location
- Type of injury
- Severity of injury
- Recurrence
- Injury mechanism (traumatic or overuse)
- Quality of the study (abbreviated STROBE scale)

Appendix 3.2. Analysis of the selected studies' methodological quality (n = 52)

Study	1	2	3	4	5	6	7	8	9	10	11	Score
Almutawa et al., 2014	+	+	+	+	-	+	+	+	+	-	+	9
Andersen et al., 2004	+	+	-	-	-	+	+	+	+	-	+	7
Arnason et al., 1996	+	+	-	-	-	+	+	+	+	-	+	7
Arnason et al., 2004	+	+	-	-	-	+	+	+	+	-	+	7
Arnason et al., 2005	+	-	-	-	-	+	+	+	+	-	+	6
Bjørneboe et al., 2010	+	+	+	-	-	+	+	+	+	-	+	8
Bjørneboe et al., 2014	+	+	+	+	-	+	+	+	+	-	+	9
Carling et al., 2010	+	+	+	-	-	+	+	+	+	-	+	8
Carling et al., 2015	+	+	+	-	-	+	+	+	+	-	+	8
Dupont et al., 2010	+	+	+	+	+	+	+	+	+	-	+	10
Dvorak et al., 2007	+	-	-	-	+	-	+	+	+	-	+	6
Dvorak, et al., 2011	+	+	+	-	-	+	+	+	+	+	+	9
Eirale et al., 2012	+	+	+	+	-	+	+	+	+	-	+	9
Eirale et al., 2013a	+	+	+	+	-	+	+	+	+	-	+	9
Eirale et al., 2013b	+	+	+	-	+	+	+	+	+	+	+	10
Ekstrand & Tropp, 1990	+	-	-	-	-	-	+	+	-	-	+	4
Ekstrand et al., 2004	+	+	-	-	-	+	+	-	+	-	+	6
Ekstrand et al., 2010	+	+	+	-	-	+	+	+	+	-	+	8
Ekstrand et al., 2011a	+	+	+	-	-	+	+	+	+	-	+	8
Ekstrand et al., 2011b	+	+	+	+	-	+	+	+	+	-	+	9
Ekstrand et al., 2013	+	+	+	+	-	+	+	+	+	-	+	9
Engström et al., 1990	+	-	-	-	-	-	+	+	+	-	+	5
Fünten et al., 2014	+	+	+	+	+	+	+	+	+	-	+	10
Gouttebarge et al., 2016	+	+	+	-	-	+	+	+	+	-	+	8
Gustafsson, 2011	+	+	+	-	+	+	+	+	+	-	+	9
Hägglund et al., 2003	+	+	+	-	-	+	+	+	-	+	+	8
Hägglund et al., 2005	+	+	-	+	-	+	+	+	+	-	+	8
Hägglund et al., 2006	+	+	+	-	-	+	+	+	+	-	+	8
Hägglund et al., 2009a	+	+	+	-	+	+	+	+	+	-	+	9
Hägglund et al., 2009b	+	+	+	-	-	+	+	+	+	-	+	8
Hägglund et al., 2016	+	+	+	-	-	+	+	+	-	-	+	7
Hassabi et al., 2010	+	-	-	-	-	+	+	+	+	-	+	6
Hawkins & Fuller, 1996	+	+	-	-	-	+	+	-	+	-	+	6
Hawkins & Fuller, 1999	+	+	-	-	+	+	+	+	+	-	+	8
Junge & Dvořák, 2013	+	-	+	-	-	+	+	+	+	-	+	7
Junge & Dvořák, 2015	+	+	+	-	-	+	+	+	+	+	+	9
Junge et al., 2004a	+	-	-	-	-	+	+	+	+	+	+	7
Junge et al., 2004b	+	+	-	-	-	+	+	+	+	-	+	7
Kristenson et al., 2013	+	+	+	+	+	+	+	+	+	-	+	10
Lee et al., 2014	+	+	+	-	-	+	+	+	+	-	-	7
Mallo et al., 2011	+	+	+	-	-	+	+	+	+	-	+	8
Nielsen & Yde, 1989	+	+	-	-	-	+	+	-	+	-	+	6

Noya et al., 2014a	+	+	+	+	-	+	+	+	+	-	+	9
Noya et al., 2014b	+	+	+	+	-	+	+	+	+	-	+	9
Pedrinelli et al., 2013	+	+	-	-	-	-	+	+	+	-	+	6
Poulsen et al., 1991	+	-	-	-	-	-	+	-	-	-	+	3
Reis et al., 2015	+	+	-	+	-	+	+	+	+	-	+	8
Shalaj et al., 2016	+	+	+	-	+	+	+	+	+	-	+	9
Stubbe et al., 2015	+	+	+	+	+	+	+	+	+	+	+	11
Waldén et al., 2005	+	+	-	+	-	+	+	+	+	-	+	8
Waldén et al., 2007	+	+	-	+	-	+	+	+	+	+	+	9
Waldén et al., 2013	+	+	+	+	-	+	+	+	+	+	+	10

The numbers of the columns corresponded to the following items of the STROBE scale.

Materials and methods

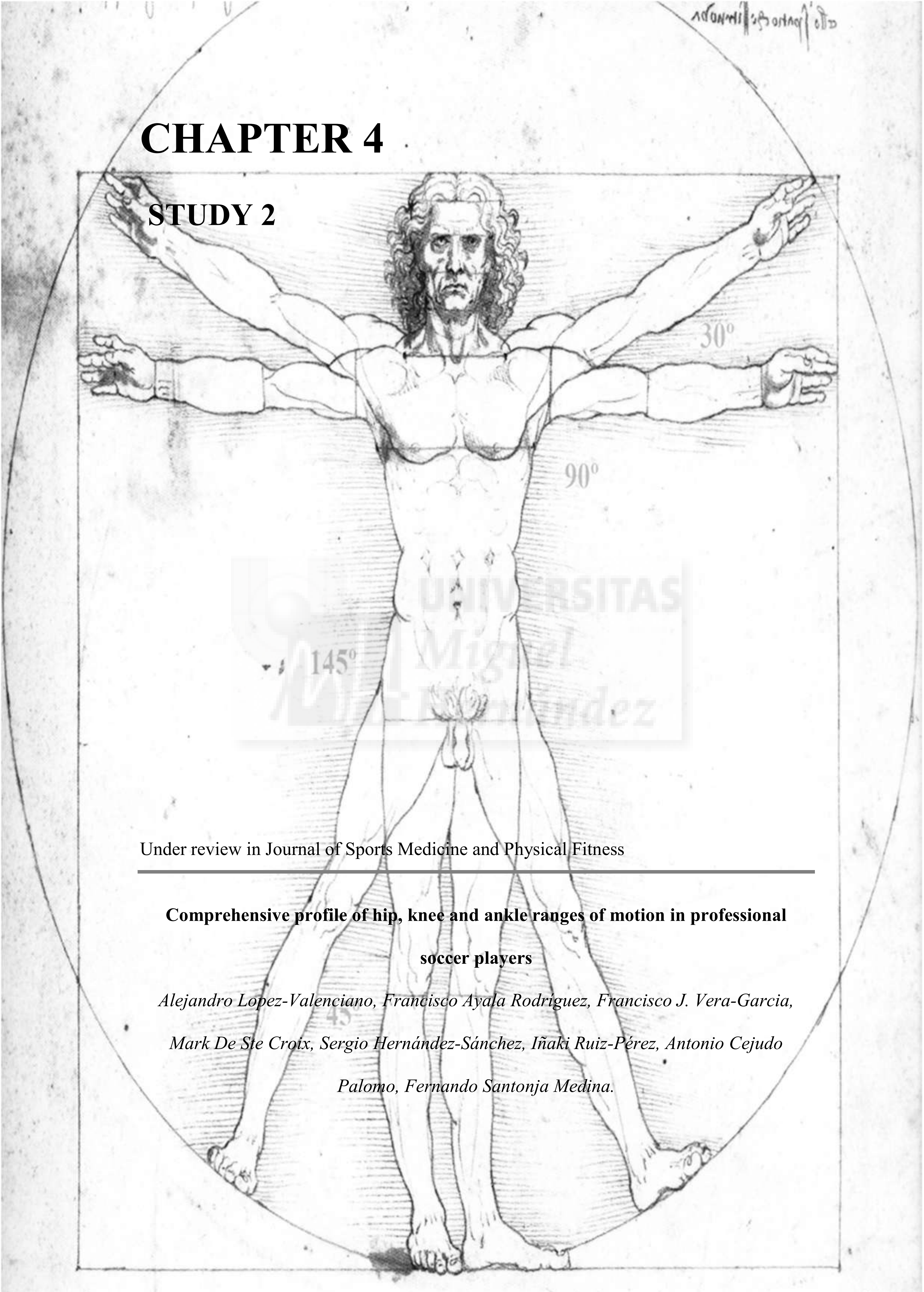
1. Describes the setting or participating locations
2. Describes relevant dates (period of recruitment, exposure, follow-up, data collection)
3. Provides statement concerning institutional review board approval and consent
4. Gives the inclusion and exclusion criteria
5. Describes injury history
6. Describes methods of follow-up

Data sources/measurement

7. Provides a definition of injury
8. Verifies injury by an independent medical professional
9. Classifies injury (severity, location and type of injury)
10. Indicates the number of participants with missing data and explain how this was addressed
11. Measures and presents exposure data

CHAPTER 4

STUDY 2



Under review in Journal of Sports Medicine and Physical Fitness

Comprehensive profile of hip, knee and ankle ranges of motion in professional soccer players

Alejandro Lopez-Valenciano, Francisco Ayala Rodríguez, Francisco J. Vera-Garcia, Mark De Ste Croix, Sergio Hernández-Sánchez, Iñaki Ruiz-Pérez, Antonio Cejudo Palomo, Fernando Santonja Medina.

CHAPTER 4

STUDY 2

Comprehensive profile of hip, knee and ankle ranges of motion in professional soccer players

by

López-Valenciano A, Ayala F, Vera-Garcia FJ, De Ste Croix M, Hernández-Sánchez S, Ruiz-Pérez I, Cejudo A, & Santonja F.

4.1. Abstract

Background: Limited ranges of motion (ROM) have been considered as a primary risk factor for some soccer injuries, but only a few studies have analysed differences in lower extremity joints. The main purposes were (a) to describe the lower extremity ROM profile in professional soccer players; and (b) to examine differences between goalkeepers and outfield players.

Methods: 82 professional male soccer players from 4 teams were measured in the 2013 preseason. Measures of passive hip (flexion with knee flexed [PHF_{KF}] and extended [PHF_{KE}], extension [PHE], abduction [PHA], external [PHER] and internal [PHIR] rotation), knee (flexion [PKF]) and ankle dorsiflexion (with knee flexed [ADF_{KF}] and extended [ADF_{KE}]) ROMs were taken. Magnitude-based inferences exploring differences between player position and legs were made.

Results: 46% of all participants showed restricted PHF_{KE} and/or around 30% showed restricted ADF_{KF} ROM values. Contrarily, most players reported normal PHF_{KF}, PHE, PHIR and PHER as well as PKF ROM scores with percentage values close to 100%. Bilateral meaningful differences for PHA, PHIR and PHER were found in approximately 30% of outfield players and goalkeepers. Statistical analysis found trivial differences between players for PHF_{KE}, PHE, PHIR, PHER, ADF_{KE} and ADF_{KF}. However, moderate differences between players were found for PHF_{KF}, PHA and PKF, with goalkeepers demonstrating higher values than outfield players.

Conclusions: The findings of this study reinforce the necessity of prescribing exercises aimed at improving PHF_{KE} and ADF_{KF} ROM within everyday soccer training routines. In addition, as some bilateral deficits were observed, unilateral training should be considered where appropriate.

Keywords: *clinical examination, injury prevention, sport therapy, muscle strain.*

4.2. Introduction

Soccer is by far the world's most popular sport, with more than 270 million participants.[1] Soccer requires players to perform many repeated high intensity movements such as sudden acceleration and deceleration, rapid changes of direction, jumping and landing tasks; as well as many situations in which players are involved in tackling to keep possession of or to win the ball.[215] The high intensity demands of movements required in soccer could lead to an overload in the joints, generating sport-specific adaptations that would cause impairments in their normal range of motion (ROM) during soccer activities and thus may result in a notable risk of injuries.[20,56,62,93,216]

Therefore, it would appear important to analyse the possible soccer-specific adaptations in the lower extremity joint ROMs at professional level in order to effectively plan and establish successful prevention and rehabilitation programmes. Some studies have analysed the impact of soccer play in some hip (flexion, extension and abduction) and knee (flexion and extension) ROMs,[20,51-56,62,93,217,218] reporting normal (compared to the sedentary population) and non-pathologic (based on the previously published cut-off scores to classify athletes at high risk of injury) ROM values. Only Daneshjoo, Rahnama, Mokhtar, & Yusof[52] have reported bilateral asymmetries (in favour of the dominant leg) in hip flexion ROM with the knee extended. These results have led some soccer health care professionals to overlook the assessment of the lower extremity joint ROMs in preseason screening sessions and to question the use of stretching exercises during both the pre- and in-season training schedules, as a preventative measure to reduce the number and impact of some soccer-related injuries.

However, when interpreting the extant literature regarding the effects of soccer play on normal lower extremity ROMs, some limitations are noted, which should be clarified before recommendations to soccer sports science and medicine practitioners can be made. For instance, it should be noted that few studies[51,52,54-56,219] have analysed whether soccer-specific adaptations would occur in the ankle and hip rotation ROMs despite the fact that restricted scores have been considered as primary risk factors for some of the most common injuries in soccer, such as ankle sprain[51,220,221] and knee osteoarthritis[222] respectively. Even less studies[20,51,223] have analysed the possible differences in lower extremity joint ROMs between goalkeepers and outfield players in order to make evidence-based training recommendations. Finally, no studies have reported whether professional players present normal or restricted hip, knee or ankle ROM values. This knowledge would allow a better understand of the possible soccer-specific adaptations in the lower extremity joint ROMs that might be caused by technical and tactical training and a lack of bilateral conditioning. Previous studies have suggested that there is a large degree of inter-player variability in ROMs[20,51,53-56,93,217,218] and thus reporting group average ROM may distort the true extent of the number of players reporting restricted ROM.

Thus, it remains to be clarified whether the repetitive loading forces generated during soccer training and match play induce alterations in the lower-extremity joint ROMs profile in professional soccer players, such as bilateral differences or as an individual deficit in one or more ROM. Furthermore, only two studies have analysed the possible differences in lower extremity ROM profiles between goalkeepers and outfield players reporting conflicting results.[20,51] Consequently, more studies are needed to address this issue, as this knowledge would allow sports science and medicine practitioners to establish specific ROM goals to be achieved by goalkeepers and outfield players through planned prevention and rehabilitation programmes.

Therefore, the aims of the present study were: (a) to describe the lower extremity ROM profile in professional soccer players; and (b) to analyse if there are ROM differences between goalkeepers and outfield players.

4.3. Methods

4.3.1. Participants

Eighty-two professional young adult male soccer players (68 outfield players and 14 goalkeepers) completed this study. Participants were recruited from 4 different soccer teams that were engaged in the professional championships of the Spanish Football Federation. Before data collection, participants completed a questionnaire containing questions about their sport-related background (player position, current level of play, dominant leg [defined as the participant´s kicking leg], sport experience); anthropometric characteristics (age, body mass, stature and body mass index); and training regimen (weekly practice frequency, hours of soccer practice per week and day, and stretching exercises and load routinely performed in their daily training sessions). Data from questionnaires reported that the sample was homogeneous in potential confounding variables, such as body mass, stature, age, training regime (one game and 4–6 days of training per week), climatic conditions, level of play (professional players), resting periods and sport experience (at least 8 years) (Table 4.1). In addition, none of the participants were involved in systematic and specific stretching regimes in the last 6 months, apart from the 1-2 sets of 15-30 s of static stretches designated for the major muscles of the lower extremities (e.g. gluteus, hamstrings, quadriceps, adductors and triceps surae) that were performed daily during their pre-exercise warm-up and post-exercise cool down phases.

The exclusion criterion was history of orthopaedic problems to the knee, thigh, hip, or lower back in the 3 months before the study because whose residual symptoms could have an impact in the habitual players´ movement competency and/or lower extremity ROM profile. The study was conducted at the end of the preseason phase of the year 2013. The time frame of the study was selected to be sure that the players recruited to each team was definitive and stable within the testing period.

Table 4.1. Demographic variables for the professional soccer players.

	Mean ± SD
Age (years)	25.5 ± 5.0
Height (cm)	180.1 ± 6.5
Body mass (kg)	75.0 ± 6.5
Years playing soccer (years)	16.1 ± 4.0
Weekly practice frequency	6.1 ± 1.2
Hours of soccer practice per week	9.8 ± 2.1
Hours of soccer practice per day	1.6 ± 0.5

SD: standard deviation.

Before any participation, experimental procedures and potential risks were fully explained to the participants in verbal and written form, and written informed consent was obtained. The experimental procedures used in this study were in accordance with the Declaration of Helsinki and were approved by the University Office for Research Ethics.

4.3.2. Testing procedure

The passive hip flexion with knee flexed (PHF_{KF}) and extended (PHF_{KE}), extension (PHE), abduction (PHA), external (PHER) and internal (PHIR) rotation; knee flexion (PKF); and ankle dorsiflexion with knee flexed (ADF_{KF}) and extended (ADF_{KE}) ROMs of the dominant and non-dominant leg were assessed following the methodology previously described[224] (Figure 4.1). These tests were selected because they have been considered appropriate by American Medical Organizations[225,226] and included in manuals of sports medicine and science[227,228] based on reliability and validity studies, anatomical knowledge, and extensive clinical and sport experience. In addition, studies from our laboratory have reported moderate to high reliability for the procedures employed (variability ranging from 4º to 9º).[224,229]

The dominant leg was defined as the participant´s preferred kicking leg. All tests were carried out by the same two physical therapists under stable environmental conditions.

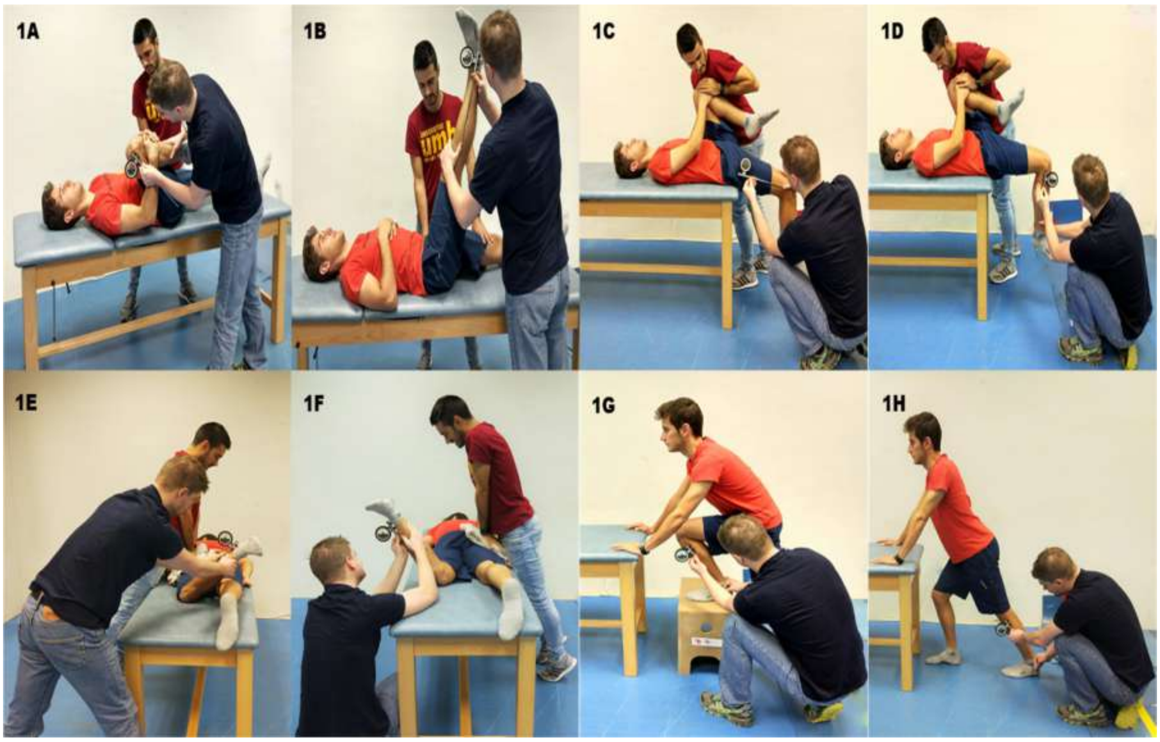


Figure 4.1. Lower extremity ranges of motion. Passive hip flexion with knee flexed test [PHF_{KF}] (1A); passive hip flexion with knee extended test [PHF_{KE}] (1B); passive hip extension [PHE] (1C); passive knee flexed [PKF]) (1D); passive hip external rotation test [PHER] (1E); passive hip internal rotation test [PHIR] (1F); ankle dorsiflexion with knee flexed test [ADF_{KF}] (1G); ankle dorsiflexion with knee extended test [ADF_{KE}] (1H).

Prior to the testing session, all participants performed the dynamic warm-up designed by Taylor, Sheppard, Lee, & Plummer.[230] The overall duration of the entire warm-up was approximately 20 min (Table 4.2). A 3-5 min rest interval between the end of the warm-up and beginning of the ROMs assessment was given to the participants because in a pilot study with 10 participants of similar age and training status, practically required some time, to get hydration and to dry their sweat prior to the ROMs assessment. More importantly, it has been shown that the effects elicited by the dynamic warm-up on muscle properties might last more than 5 min[231] and hence, decreases in ROM values within the 3-5 min rest interval were not expected.

Table 4.2. Pre-assessment dynamic warm up*.

Exercise	Duration
1. High knees	3 set over 20 m
2. Butt flicks	3 set over 20 m
3. Carioca	3 set over 20 m each side
4. Dynamic hamstring swings	10 repetitions each leg
5. Dynamic groin swings	10 repetitions each leg
6. Arm swings: forwards and backwards	10 repetitions each direction
7. Faster high knees (shorter stride)	4 sets over 10 m
8. Swerving	2 sets over 30 m at 70% of maximum pace
9. Side stepping	2 sets over 30 m at 80% of maximum pace
10. Spiderman walks	1 set over 20 m
11. Sideways low squat walks	1 set x 10 steps each direction
12. Upper body rotations	10 repetitions each leg
13. Vertical jump	5 repetitions building in intensity
14. Run through	– 2 sets x 20 m at 70% of maximum pace – 2 sets x 20 m at 80% of maximum pace – 1 set x 20 m at 90% of maximum pace
15. Countermovement jump then 5 m sprint	– 2 sets x 5 m at 90% of maximum pace – 1 sets x 5 m at 95% of maximum pace
16. Sprint for 5 m then countermovement jump	2 sets x 5 m

m: meters; *: warm up programme extracted from Taylor et al.[230]

After the warm-up, participants were instructed to perform, in a randomised order, two maximal trials of each ROM test for each leg, and the mean score for each test was used in the analyses. Participants were examined wearing sports clothes and without shoes. A 30 s rest was given between trials, legs and tests.

One or both of the following criteria determined the endpoint for each test: (a) palpable onset of pelvic rotation, and/or (b) the participant feeling a strong but tolerable stretch, slightly before the occurrence of pain.

4.3.3. Statistical analysis

Prior to the statistical analysis, the distribution of raw data sets was checked using the Kolomogorov-Smirnov test and demonstrated that all data had a normal distribution ($p > 0.05$). Descriptive statistics including means and standard deviations were calculated for hip, knee and ankle ROM measures separately by player position (outfield players and goalkeepers) and leg (dominant and non-dominant).

Furthermore, in each participant, the hip, knee and ankle ROM scores were categorized as normal or restricted according to the reference values previously reported to consider an athlete as being more prone to suffer an injury.[232-236] In cases where no cut-off scores for detecting athletes at high risk of injury had been previously reported (e.g. PHA and PHIR ROM), they were compared with data generated on the general population. Thus, ROM values were reported as restricted according to the following cut-off scores: < 114º for the PHF_{KE} ROM,[232] < 80º for the PHF_{KF} ROM,[233] < 50º for the PHA ROM,[237] < 25º for the PHIR ROM,[235] < 25º for the PHER ROM,[238] <0º for the PHE ROM,[236] < 17º for the ADF_{KE} ROM,[239] and < 34º ADF_{KF}ROM.[234]

To make comparisons with the results reported in previous similar studies, magnitude-based inferences on differences between player position (outfield players versus goalkeepers) and leg (dominant versus non-dominant) were determined using a spreadsheet designed by Hopkins[173] for change scores between paired comparisons for each ROM variable. This analysis determines the chances that the differences are substantial or trivial when a value for the smallest worthwhile change is entered. The cut off score of > 6º proposed by Fousekis, Tsepis, Poulmedis, Athanasopoulos, & Vagenas[93] determined the smallest substantial/worthwhile change for both the inter-player and leg comparisons for each of the ROM variables. The qualitative descriptors proposed by Hopkins[240] were used to interpret the probabilities that the true affects are harmful, trivial or beneficial: < 1%, almost certainly not; 1–4%, very unlikely; 5– 24%, unlikely or probably not; 25–74%, possibly or may be; 75–94%, likely or probably; 95–99%, very likely; > 99%, almost certainly. Effect sizes, which are standardised values that permit the determination of the magnitude of differences between groups or experimental conditions were also calculated for each of the variables using the method and descriptors previously described by Cohen.[241] Based on Fousekis et al.,[93] the number of players with side-to-side differences (> 6º) in each ROM measure were also calculated.

Analysis was completed using SPSS version 20 (SPSS Inc, Chicago, IL, USA) and an online spreadsheet (www.sportsci.org).

4.4. Results

Tables 4.3 and 4.4 show the descriptive ROM values (mean ± SD) for passive hip (PHF_{KF}, PHF_{KE}, PHE, PHA, PHIR and PHER), knee (PKF) and ankle (ADF_{KE} and ADF_{KF}) for both, outfield players and goalkeepers, respectively.

Statistical analysis reported no meaningful differences (trivial effect with a probability > 99%) between dominant and non-dominant legs for each ROM variable in both outfield players (Table 4.3) and goalkeepers (Table 4.4).

Statistical analysis also reported trivial differences (trivial effect with a probability of 84-100%; $d < 0.2$) between players (outfield players and goalkeepers) for PHF_{KE}, PHE, PHIR, PHER, ADF_{KE} and ADF_{KF} ROM measures (Table 4.5). However, moderate differences (possibly meaningful effect with a probability of 62-71%; $d >0.40$) between players were found for PHF_{KF}, PHA and PKF, with goalkeepers showing higher scores than outfield players.

Table 4.3. Field based players´ descriptive values and inference about side-to-side difference for hip (flexion, extension, abduction, internal and external rotation), knee (flexion) and ankle (dorsiflexion with knee flexed and extended) ranges of motion (n = 68).

Range of motion (º)	Dominant leg		No-dominant leg		Players with bilateral differences > 6º	Inference[a,b]
	Mean ± SD	Qualitative Outcome*	Mean ± SD	Qualitative Outcome*		
PHF_{KF}	145.9 ± 8.1	Normal (0)	147.3 ± 7.6	Normal (0)	6	Most likely trivial (0/100/0)
PHF_{KE}	80.3 ± 10.9	Normal (28)	81.1 ± 11.3	Normal (26)	8	Most likely trivial (0/100/0)
PHA	63.3 ± 9.1	Normal (6)	60.6 ± 8.2	Normal (6)	20	Most likely trivial (0/100/0)
PHIR	47.1 ± 8.0	Normal (1)	45.3 ± 7.9	Normal (0)	16	Most likely trivial (0/100/0)
PHER	49.9 ± 9.8	Normal (1)	50.7 ± 9.8	Normal (0)	22	Most likely trivial (0/100/0)
PHE	8.9 ± 8.8	Normal (11)	9.8 ± 8.5	Normal (10)	4	Most likely trivial (0/100/0)
PKF	126.9 ± 13.6	Normal (0)	124.6 ± 13.5	Normal (0)	14	Most likely trivial (0/100/0)
ADF_{KE}	36.1 ± 5.7	Normal (0)	36.3 ± 5.7	Normal (0)	5	Most likely trivial (0/100/0)
ADF_{KF}	37.2 ± 6.6	Normal (21)	37.8 ± 6.1	Normal (18)	5	Most likely trivial (0/100/0)

PHF_{KF}: passive hip flexion with knee flexed test; PHF_{KE}: passive hip flexion with knee extended test; PHA: passive hip abduction test; PHIR: passive hip internal rotation test; PHER: passive hip external rotation test; PHE: passive hip extension test; PKF: passive knee flexion test; ADF_{KE}: ankle dorsiflexion with knee extended test; ADF_{KF}: ankle dorsiflexion with knee flexed test.

º: degrees; *: qualitative score of the mean range of motion, in parentheses the number of players with a restricted range of motion score according to previously published cut-off scores (see Statistical analysis section).

[a] Substantial is an absolute change in performance of > 6º for all ROM measures for passing accuracy (see Methods).

[b] If chance of benefit and harm both > 5%, true effect was assessed as unclear (could be beneficial or harmful). Otherwise, chances of benefit or harm were assessed as follows: < 1%, almost certainly not; 1-5%, very unlikely; > 5-25%, unlikely; > 25-75%, possible; > 75-95%, likely; > 95-99%, very likely; > 99%, almost certain.

Table 4.4. Goalkeepers´ descriptive values and inference about side-to-side difference for hip (flexion, extension, abduction, internal and external rotation), knee (flexion) and ankle (dorsiflexion with knee flexed and extended) ranges of motion (n = 14).

Range of motion (º)	Dominant leg		No-dominant leg		Players with bilateral differences > 6º	Inference[a,b]
	Mean ± SD	Qualitative Outcome*	Mean ± SD	Qualitative Outcome*		
PHF_{KF}	150.9 ± 9.4	Normal (0)	151.8 ± 7.2	Normal (0)	0	Most likely trivial (0/100/0)
PHF_{KE}	80.3 ± 10.1	Normal (7)	79.5 ± 10.7	Restricted (8)	2	Most likely trivial (0/100/0)
PHA	67.9 ± 7.6	Normal (0)	66.6 ± 9.8	Normal (1)	4	Most likely trivial (0/100/0)
PHIR	49.4 ± 10.5	Normal (0)	47.9 ± 6.3	Normal (0)	5	Most likely trivial (0/100/0)
PHER	50.8 ± 7.6	Normal (0)	48.5 ± 8.3	Normal (0)	4	Most likely trivial (0/100/0)
PHE	12.2 ± 7.4	Normal (0)	12.7 ± 7.8	Normal (0)	1	Most likely trivial (0/100/0)
PKF	131.7 ± 10.9	Normal (0)	131.4 ± 13.2	Normal (0)	3	Most likely trivial (0/100/0)
ADF_{KE}	36.6 ± 5.1	Normal (0)	37.0 ± 5.1	Normal (0)	3	Most likely trivial (0/100/0)
ADF_{KF}	37.5 ± 7.1	Normal (2)	40.6 ± 4.7	Normal (2)	2	Most likely trivial (0/100/0)

PHF_{KF}: passive hip flexion with knee flexed test; PHF_{KE}: passive hip flexion with knee extended test; PHA: passive hip abduction test; PHIR: passive hip internal rotation test; PHER: passive hip external rotation test; PHE: passive hip extension test; PKF: passive knee flexion test; ADF_{KE}: ankle dorsiflexion with knee extended test; ADF_{KF}: ankle dorsiflexion with knee flexed test.

º: degrees; *: qualitative score of the mean range of motion, in parentheses the number of players with a restricted range of motion score according to previously published cut-off scores (see Statistical analysis section).

[a] Substantial is an absolute change in performance of > 6º for all ROM measures for passing accuracy (see Methods).

[b] If chance of benefit and harm both > 5%, true effect was assessed as unclear (could be beneficial or harmful). Otherwise, chances of benefit or harm were assessed as follows: < 1%, almost certainly not; 1-5%, very unlikely; > 5-25%, unlikely; > 25-75%, possible; > 75-95%, likely; > 95-99%, very likely; > 99%, almost certain.

Table 4.5. Inter-group differences (outfield players vs. goalkeepers) for passive hip flexion with knee flexed [PHF_{KF}] and extended [PHF_{KE}], extension [PHE], abduction [PHA] and rotation (external [PHER] and internal [PHIR]), knee (flexion [PKF]) and ankle (dorsiflexion with knee flexed [ADF_{KF}] and extended [ADF_{KE}]) range of motion values (dominant leg). Chances that the true effects were substantial and practical assessments of the effects are also shown.

Range of motion (°)	Change^{T}	Effect Size (d)	Chances that the true effects[a] were positive/trivial/negative			Qualitative inference[b]
HF_{KF}	-5.0 (-10.4 to 0.4)	-0.49	0	63	37	Possibly meaningful
PHF_{KE}	0.0 (-5.8 to 5.8)	0.00	5	91	4	Likely trivial
PHA	-4.6 (-9.1 to -0.1)	-0.56	0	71	29	Possibly meaningful
PHIR	-2.3 (-8.3 to 3.7)	-0.20	1	84	14	Likely trivial
PHER	-0.9 (-5.4 to 3.6)	-0.11	1	96	3	Very likely trivial
PHE	-3.2 (-7.6 to 1.2)	-0.40	0	86	14	Likely trivial
PKF	-4.8 (-11.2 to 1.7)	-0.40	1	62	37	Possibly meaningful
ADF_{KE}	-0.5 (-3.5 to 2.4)	-0.10	0	100	0	Most likely trivial
ADF_{KF}	-0.4 (-4.4 to 3.7)	-0.05	1	98	1	Very likely trivial

°: degrees; T: mean ± 90% confidence limits.

[a] Substantial is an absolute change in performance of > 6° for all range of motion measures for passing accuracy (see Methods).

[b] If chance of benefit and harm both > 5%, true effect was assessed as unclear (could be beneficial or harmful). Otherwise, chances of benefit or harm were assessed as follows: < 1%, almost certainly not; 1-5%, very unlikely; > 5-25%, unlikely; > 25-75%, possible; > 75-95%, likely; > 95-99%, very likely; > 99%, almost certain.

4.5. Discussion

The main findings of this study reported average values classified as normal (based on the reference values reported in previous studies) for passive hip (flexion, extension, abduction and rotation), knee (flexion) and ankle (dorsiflexion) ROMs for both outfield players and goalkeepers. Similar results have been found in previous studies[20,51-53,55,56,62,93,217] that have described the lower extremity ROM profile of soccer players. From this standpoint, no specific adaptations in the lower extremity joint ROMs would be expected as a consequence of soccer training and match play at professional levels and hence, no further injury prevention measures need to be considered, which are aimed at improving ROMs.

However, when a novel and more comprehensive analysis is carried out, the current data indicates that a large number of the soccer players demonstrate restricted PHF_{KE} (cut-off score < 80°; outfield players ≈ 40%; goalkeepers ≈ 50%)[233] and/or ADF_{KF} (cut-off score < 34°; outfield players ≈ 30%; goalkeepers ≈ 28%)[234] ROM values. These latter results are in conflict with the findings reported by previous studies that have described the lower extremity ROM profile of soccer players using average ROM scores.[20,51-53,55,56,62,93,217] This discrepancy might be explained by the fact that the average PHF_{KE} and ADF_{KF}ROM values, although categorized as normal, are close to the restricted cut-off score previously published (80º and 34º respectively)[233,234] if the inter-player variability is not taken into account the findings might be biased. As a consequence, these biased results might cause an unrealistic diagnostic of non-soccer-specific adaptations in the lower extremity joint ROMs. Comparisons with other previously published findings are not possible as there appears to be no previous study analysing the ROM of hip, knee and ankle using the same comprehensive analysis carried out in the current study.

The large percentage of players reporting restricted PHF_{KE} and ADF_{KF} ROM in the current study might be explained by the demands of soccer training and match play that requires players to perform many repeated high intensity movements such as sudden acceleration and deceleration, rapid changes of directions, jumping and landing tasks. These movements impose strong concentric and eccentric loads on the hip flexor and ankle dorsiflexion muscles (posterior kinetic chain) at shortened contracted positions.[242-244] When these actions are repeated several times during training sessions and games, they have the potential to generate muscle damage that without the proper recovery and protective measures, they might induce impairments in the mechanical and neural properties of the muscle-tendon units, including a reduction in their normal ROM and strength loss.[245]

In addition, another factor that might have contributed to these restricted ROM values could be the demanding competitive calendar of players at professional levels that can result in athletes focusing on competition and thus compromising training, leading to sub-optimal recovery and preparation. These deficits have been suggested as predisposing factors for increasing the likelihood of some of the most prevalent hip and knee pathologies in soccer players such as hamstring muscles strains,[51,56,62,92,93] patellar tendinopathy[56,246] and ankle sprain.[51,220,221] Based on the present results, sports science and medicine practitioners should include during both, the pre- and in-season training schedules, stretching exercises of the hip, enhancing hip flexion ROM with the knee extended; and ankle, enhancing dorsiflexion ROM with the knee flexed. It seems important to suggest that coaches and strength and conditioning specialists should educate the players in order to be able to distinguishing between the stretching routines used for improving joint ROM (e.g. static and proprioceptive neuromuscular facilitation stretching routines during the training sessions) and the one used as part of the warm-up process (e.g. dynamic stretching exercises), targeting to activate the muscle groups involved in a specific performance task.[247] Therefore, and based on the documented acute negative effect of static stretching on maximal muscle performance,[248] routines aimed at improving ROM values that usually include static stretching exercises should be performed at the end of the training sessions or even better as separate training sessions.

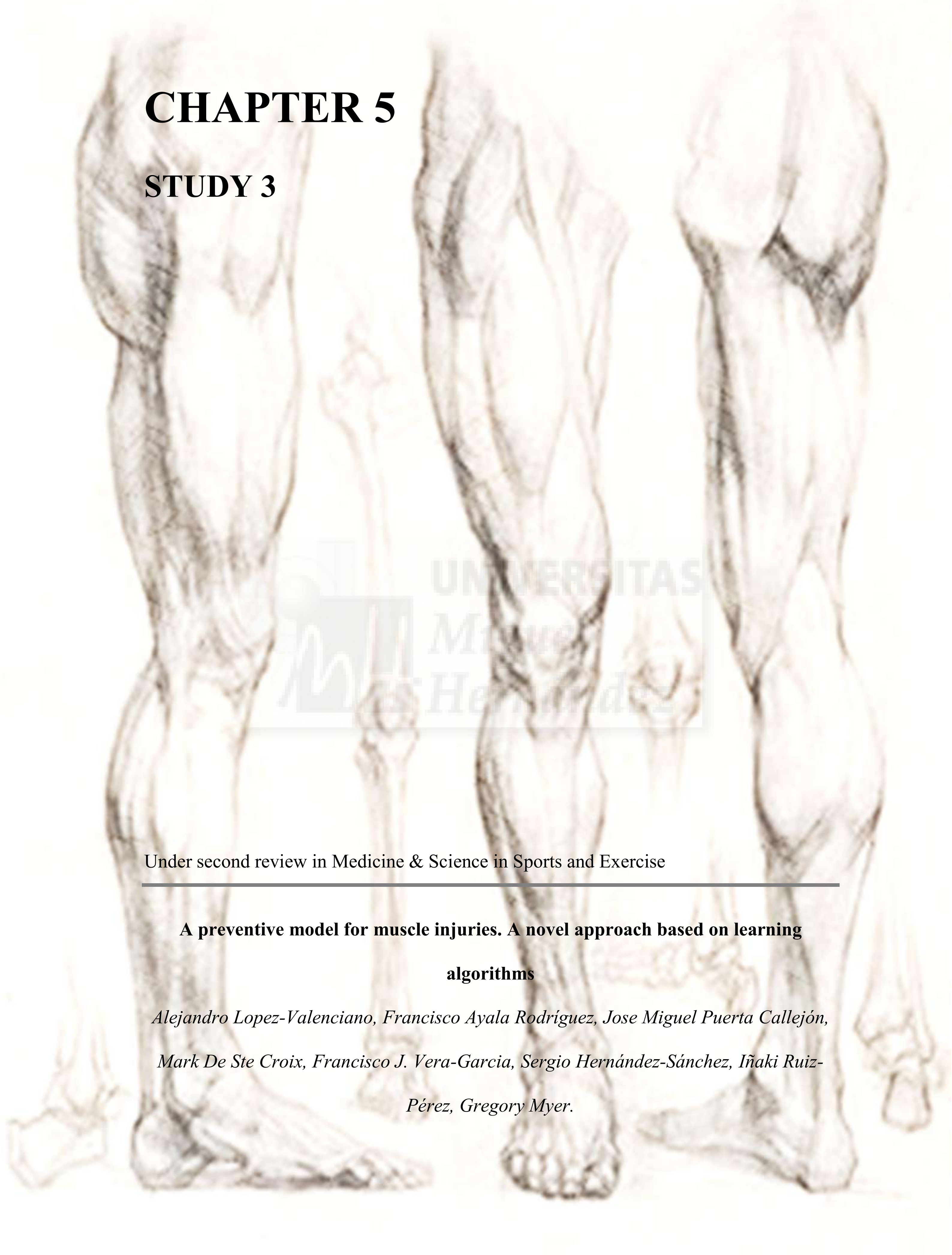
The results of the current study also found non-clinically relevant bilateral differences (> 6º) between the dominant and non-dominant leg joint ROM average values in both outfield players and goalkeepers. However, by calculating the number of players with bilateral differences greater than 6º in any hip, knee and ankle ROM measure, approximately 30% of the players (outfield players and goalkeepers) were identified for PHA, PHIR and PHER. In particular, the bilateral differences for PHA and PHIR reported were mostly in favour of the dominant leg for the outfield players (16 up to 20 cases and 13 up to 16 cases for PHA and PHIR ROMs respectively). The asymmetrical and repeated technical gestures of kicking and controlling the ball using mainly the dominant leg might be a plausible explanation for the bilateral differences in favour of the dominant leg, identified in the current study. Thus, the backswing phase of kicking (e.g. volley) and controlling the ball may reflect in some cases a dynamic stretching for the hip external and adductor muscles which may increase the hip internal and abduction ROMs respectively. In addition, and similar to what has been found in tennis players,[249] the higher number of repetitive and powerful internal rotational movements generated in the stance leg (non-dominant) during the technical gesture of kicking (forward swing) to transfer power to the final part of the movement could lead to microtrauma and capsular contracture, causing a hip internal rotation ROM deficit in many of the male players. Conversely, there was not a clear pattern for PHER ROM so that almost the same number of outfield players with bilateral differences reported greater values in the dominant and non-dominant leg. An explanation for this discrepancy has not been found.

The same circumstance was found in the goalkeepers so there appears not to be clear patterns for any meaningful bilateral difference found for PHA as well as PHIR and PHER ROM measures. Perhaps, the small sample size of goalkeepers (n = 14) might explain why we did not observe any pattern. Although still inconclusive, some studies have suggested that bilateral asymmetries of lower extremity ROMs may alter the kinetic patterns of lower extremity function during the production of excessive and asymmetrical forces in explosive sports activities, such as kicking and cutting in soccer and this might play a role in the mechanisms that predispose a soccer player to suffer an injury (mainly muscle strains).[93,250] The current study also identified the presence of moderate differences (possibly meaningful effect with a probability of 62-71%; d > 0.40) between players for PHF_{KF}, PHA and PKF ROM measures, with goalkeepers showing higher values than outfield players. Similar PKF ROM differences in favour of goalkeepers were found by Arnason, Sigurdsson, Gudmundsson, Holme, Engebretsen, & Bahr.[20] However, Bradley, & Portas[51] found differences in PHF_{KF}, PHA and PKF ROM measures between outfield players and goalkeepers. Perhaps, the higher ROM scores shown by goalkeepers may be due to their specific physical demands as they need greater ROM values to cover a large perimeter of the goal and to stretch as much as possible to save or deflect shots.[20]

Some limitations to the study must be acknowledged. The age distribution of participants was relatively narrow and the goalkeepers' sample size was small. Moreover, the use of different testing methodologies (e.g. active ROMs) makes comparisons difficult.

4.6. Practical applications

The findings of this study reinforce the necessity of prescribing exercises aimed at improving PHF_{KE} and ADF_{KF} ROM values in the everyday soccer training routines of professional male players. Furthermore, the findings of this study also indicate no significant differences (< 5º) in ROM for the hip, knee and ankle between outfield players and goalkeepers and hence, exercises designed and prescribed in applied settings do not have to be adapted for individuals and could be delivered as group exercise. Although we found few ROM deficits in the current sample, some bilateral differences were observed and unilateral training should be considered in sports where training might promote bilateral differences. This is especially so in professional soccer were repetitive movements are undertaken that involve a kicking and stance leg which develop bilateral deficits.



CHAPTER 5

STUDY 3

Under second review in Medicine & Science in Sports and Exercise

---

A preventive model for muscle injuries. A novel approach based on learning algorithms

Alejandro Lopez-Valenciano, Francisco Ayala Rodríguez, Jose Miguel Puerta Callejón, Mark De Ste Croix, Francisco J. Vera-Garcia, Sergio Hernández-Sánchez, Iñaki Ruiz-Pérez, Gregory Myer.

CHAPTER 5

STUDY 3

A preventive model for muscle injuries. A novel approach based on learning algorithms

by

Lopez-Valenciano A, Ayala F, Puerta JM, De Ste Croix M, Vera-Garcia FJ, Hernández-Sánchez S, Ruiz-Pérez I, & Myer G.

5.1. Abstract

Background: The application of contemporary statistical approaches coming from machine learning and data mining environments, which build more robust predictive models to identify players at high risk of injury, might support injury prevention strategies of the future.

Purpose: The purpose was to analyse and compare the behaviour of some machine learning methods in order to select the best performing injury risk factor model to identify players at risk of lower extremity muscle injuries (MUS_{INJ}).

Study Design: Prospective cohort study.

Methods: A total of 132 male professional soccer and handball players underwent a preseason screening evaluation of a number of personal, psychological and neuromuscular measures. Furthermore, injury surveillance was employed to capture all the MUS_{INJ} occurring in the 2013/2014 season. The predictive ability of several models built by applying a range of learning techniques were analysed and compared.

Results: There were 32 MUS_{INJ} over the follow up period, 21 (65.6%) of which corresponded to the hamstrings, four to the adductors (12.5%), four to the triceps surae (12.5%) and three to the quadriceps (9.3%). A total of 13 injures occurred during training and 19 during match. Three players were injured twice during the observation period so the first injury was used leaving 29 MUS_{INJ} that were used to develop the predictive models. The model generated by the SmooteBoost technique with a cost-sensitive alternating decision tree as base classifier reported the best evaluation criteria (area under the receiver operating characteristic curve score = 0.747) and hence was considered the best for predicting MUS_{INJ}.

Conclusions: The prediction model showed high accuracy for identifying professional soccer and handball players at risk of MUS_{INJ}. Therefore, the model developed might help in the decision-making process for injury prevention.

Keywords: *injury prevention, machine learning techniques, modelling, screening, soccer.*

5.2. Introduction

Lower extremity muscle injuries (MUS_{INJ}) are very common in professional sports, such as soccer,[41] rugby[251] and handball.[252] These sports require sudden acceleration and deceleration tasks with rapid changes of directions,[118] as well as many situations in which players are required to repetitively kick a ball[253] and/or to be involved in tackling to keep possession of or to win the ball.[254] Professional soccer teams with a 25-player squad could expect 15 MUS_{INJ} each season and MUS_{INJ} can account for more than a quarter of all lost time from injuries.[41] In particular, injuries to four major muscle groups of the lower extremity (e.g. adductors, hamstrings, quadriceps, and calf) comprise more than 90% of all MUS_{INJ} in soccer.[41] Therefore, there is a clear necessity to develop and implement strategies aimed at preventing and reducing the number and severity of MUS_{INJ} in professional players.

Prior to establishing MUS_{INJ} prevention programmes, it is essential to identify players at high risk of MUS_{INJ} through a validated screening programme.[7] Bahr[7] in a recently published thought-provoking critical review, suggested that prior to consider a screening programme as valid to predict and prevent sports injuries it should have successfully overcome three steps. The first step is to identify those potential risk factors that have demonstrated a strong relationship with injury in prospective studies and then define appropriate cut-off values. The second step is to determine the validity of the screening tests used to measure the risk factors to predict new injuries in a new athlete population. Finally, in the third step studies should document that an intervention programme targeting athletes identified as being at high risk using the developed screen must be more beneficial than the same intervention programme given to all athletes.

In recent years, a substantive effort has been made by the scientific community and medical practitioners to identify strong risk factors associated with the occurrence of muscle injuries. Thus, some prospective studies, but not all, have identified previous injury,[20,23,57] older age,[20,23,53] poor flexibility,[20,53,93] fatigue[24] and decreased muscle strength or strength imbalances[57,93,118] as potential risk factors associated with MUS_{INJ}. Despite the fact that significant associations (causal relationship) were found between these risk factors and MUS_{INJ}, the ability of the cut-off scores proposed to predict injuries are not acceptable for screening purposes. In particular, most of the cut-off scores reported in previous studies show good true negative rates (e.g. how many individuals with a negative score were not injured), however the true positive rates were very low (e.g. how many individuals with a positive score were injured). The consequence of this has led Bahr[7] to conclude that: a) find statistically significant association between a test result and MUS_{INJ} is not sufficient evidence to use the test to predict who is at risk of injury; and b) there is no screening test available to predict sports injuries (including MUS_{INJ}) with adequate properties and consequently the exercises included in intervention programmes are not evidence-based supported as the link between risk factors and injury incidence remain to be established. Furthermore, these two conclusions appear to be supported by the fact that recent evidence has demonstrated that $\mathrm{MUS_{INJ}}$ incidence has not only decreased, but it has increased slightly throughout the last years.[48]

Perhaps one the main reasons behind the lack of available valid screening programmes to predict players at high risk of suffering a sport injury, including $\mathrm{MUS_{INJ}}$, could be based on the use of traditional and/or sub optimal statistical approaches (e.g. multivariate logistic regression analysis). Logistic regression models generally do not deal well with class imbalance problems, such as the $\mathrm{MUS_{INJ}}$ phenomenon, in which the number of injured players (minority class) prospectively reported is always much lower than the non-injured players (majority class).[113] Thus, in many scenarios including $\mathrm{MUS_{INJ}}$, traditional multivariate analyses are often biased (for many reason) towards the majority class (known as the ‘‘negative’’ class) and therefore, there is a higher misclassification rate for the minority class instances (called the ‘‘positive’’ examples), which represent the most important concept.[114] Furthermore, another limitation of the current body of the literature is based on the fact that most of the predictive models available in the sports medicine context might be over-fitted (i.e. their predictive ability is adjusted to the data set used in their building process). In this sense, no study (to the authors´ knowledge) has carried out any type of cross-validation process to analyse the predictive properties of the models developed in a cohort of players different from those used for building them. Another reason for the limited validation screening programmes might be due to the fact that the previously mentioned studies have analysed the predictive ability of each risk factor in isolation or in conjunction with just two or three risk factors. However, the $\mathrm{MUS_{INJ}}$ phenomenon has been considered as being multifactorial, in which several factors have an influence on it, and in some cases interact among them.[126] Therefore, it might be possible that the individual ability of each potential risk factor to impact on the likelihood to suffer a $\mathrm{MUS_{INJ}}$ could be very small and in most cases not statistically significant unless analysed in conjunction with other known factors simultaneously as a complex component or factor.

The application of contemporary statistical approaches (e.g. supervised learning algorithms) coming from machine learning and data mining environments that have been specifically designed to deal with class imbalance problems[113] and that can manage a large number of variables in order to develop a robust predictive model might shed light on this problematic in the sport medicine setting. In fact, these statistical approaches have been applied, among others, in several medical diagnosis studies reporting excellent results.[136]

Therefore, the main purpose of the current prospective study was to analyse and compare the behaviour of some learning methods in order to select the best performing injury risk factor model to predict $\mathrm{MUS_{INJ}}$ in a cohort of professional players.

5.3. Methods

5.3.1. Participants

A total of 132 male professional soccer (n = 98) and handball (n = 34) players took part in the current study. Soccer players were recruited from four different soccer teams that were engaged in the 1st (one team, n = 25) and 2nd B (three teams, n = 73) Spanish National Soccer League divisions. Handball players were recruited from three different handball teams that were engaged in the 1st (one team, n = 11) and 3rd (two teams, n = 23) National Handball League divisions. The sample was homogeneous in potential confounding variables, such as body mass, stature, age, training regime (one game and 4–6 days of training per week), climatic conditions, level of play, resting periods and sport experience (at least 8 years).

The exclusion criteria were: a) presence of orthopaedic problems that prevented the proper execution of one or more of the neuromuscular tests selected for this study; and b) players who were transferred to other clubs and did not finish the 9-month follow up period. Only new injuries were used for any player sustaining multiple $MUS_{INJ.}$

Prior to study participation, experimental procedures and potential risks were fully explained to the participants in verbal and written form, and written informed consent was obtained from them. An Institutional Research Ethics committee approved the study protocol prior to data collection, conforming to the recommendations of the Declaration of Helsinki.

5.3.2. Study design

A prospective cohort design was used to address the purposes of this study. In particular, all the MUS_{INJ} accounted for within the 9 months (2013/2014 season) following the initial testing session were prospectively collected for all players.

Players underwent a preseason evaluation of a number of personal, psychological and neuromuscular measures, most of them considered potential sport-related injury risk factors. For each soccer and handball team, the testing session was conducted at the preseason phase of the year.

5.3.3. Testing procedure

The testing session had a total duration of approximately 120 min and was divided into three different parts (Figure 5.1). The first part of the test session was used to obtain information related to the participants' personal or individual characteristics (5 min). The second part was designed to assess psychological measures related to sleep quality and athlete burnout (10 min). Finally, the third part of the session was used to assess a number of neuromuscular measures (105 min).

Each of the 8 testers who took part in this study conducted the same tests throughout all the testing sessions and they were blinded to the purposes of this study. All testers had more than 4 years of experience in neuromuscular assessment.

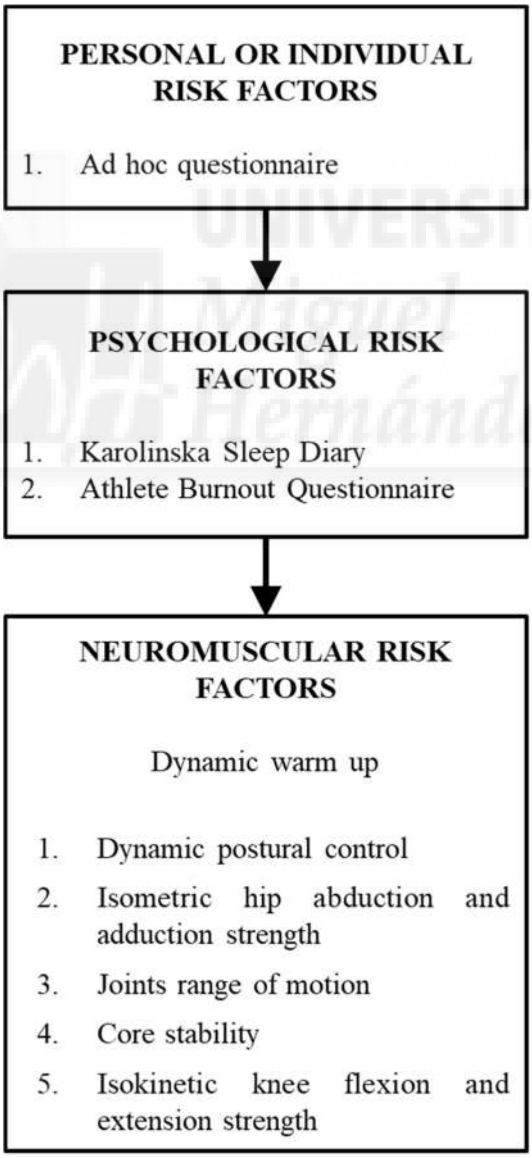


Figure 5.1. Graphical representation of testing procedure.

5.3.3.1. Personal or individual risk factors

The ad hoc questionnaire designed by Olmedilla, Laguna, & Redondo[255] was used to record personal or individual features that have been defined as potential non-modifiable risk factors for sport injuries. Through this questionnaire sport-related background (sport, player position, current level of play, dominant leg [defined as the participant´s kicking leg]) and demographic (age, body mass, stature and body mass index) features were recorded. In addition, the presence within the last season (yes or no) of MUS_{INJ} with a total time taken to resume full training and match > 8 days was also recorded (self-reported). Appendix 5.1 displays a description of all the personal risk factors recorded.

5.3.3.2. Psychological risk factors

Sleep quality and athlete burnout variables were measured through two validated and worldwide used likert scales. The Spanish version of the Pittsburgh Sleep Diary[256] was used to measure the sleep quality of the soccer and handball players. The final score of this scale was determined as the average of the scores obtained in each of its 7 items. The Spanish version of the Athlete Burnout Questionnaire[257] was used to assess the three different dimensions that comprise athlete burnout: a) physical/emotional exhaustion; b) reduced sense of accomplishment; and c) sport devaluation. Specifically, it is a likert scale comprising 15 items, 5 per factor, which employs a response format in ordered categories, with five alternatives: almost never (1), not very often (2), sometimes (3), often (4) and almost always (5). Appendix 5.2 displays a description of all the psychological risk factors recorded.

5.3.3.3. Neuromuscular risk factors

Prior to the neuromuscular risk factor assessment, all participants performed the dynamic warm-up designed by Taylor et al.[230] This warm-up routine was chosen because it reflects the standard warm-up structure (aerobic exercises + dynamic stretching exercises + sport-specific movements executed at, or just below game intensity) that might be the most widely used in soccer and handball. In addition, the effects elicited by this dynamic warm-up routine have been demonstrated to be enough to optimise the subsequent physical performance in elite athletes.[230] The overall duration of the entire warm-up was approximately 15-20 min. The assessment of the neuromuscular risk factors started 3-5 min after the dynamic warm-up.

In the experimental session, participants were assessed from a number of neuromuscular performance measures obtained from five different testing manoeuvres: 1) dynamic postural control; 2) isometric hip abduction and adduction strength; 3) lower extremity joint ranges of motion; 4) core stability; and 5) isokinetic knee flexion and extension strength.

The order of the tests was consistent for all participants and was established with the intention of minimizing any possible negative influence among variables. A 5-min rest interval was given between consecutive testing manoeuvres.

5.3.3.4. Dynamic postural control

Dynamic postural control was evaluated using the Y-Balance device® and following the guidelines described by Shaffer et al.[258]

The distance reached in each direction (anterior, posteromedial and posterolateral) was normalised by dividing by the previously measured leg length to standardize the maximum reach distance ([excursion distance/leg length] x100 = % maximum reach distance).[258] The bilateral ratio (dominant/non-dominant score) of each direction was also calculated. A bilateral ratio higher than 10% was considered as asymmetry. Finally, to obtain a global measure of the balance test for each leg, data from each direction were averaged to calculate a composite score.

5.3.3.5. Isometric hip abduction and adduction strength

Isometric hip abduction and adduction peak torques of the dominant and non-dominant leg were assessed with a portable handheld dynamometer (Nicholas Manual Muscle Tester, Lafayette Indiana Instruments) in a supine lying position on a test bench with the participants' legs extended and following the methodology described by Thorborg, Petersen, Magnusson, & Hölmich.[259] Briefly, participants performed five trials of 5-second isometric maximal voluntary contraction for each hip movement. The mean of the three most closely related trials were used for the subsequent statistical analyses. Unilateral hip abductor/adductor peak torque ratio defined as the hip adductor peak torque divided by hip abductor peak torque was calculated for each leg. Furthermore, the hip abduction and adduction bilateral ratios were also determined as the quotient of the dominant hip mean isometric peak value by the non-dominant hip mean isometric peak value. A side-to-side difference higher than 10% was defined as bilateral asymmetry.

5.3.3.6. Lower extremity joint ranges of motion (ROM)

The passive hip flexion with knee flexed and extended, extension, abduction, external and internal rotation, knee flexion and ankle dorsiflexion with knee flexed and extended ROMs of the dominant and non-dominant legs were assessed following the methodology previously described.[224] Furthermore, for each joint ROM measure, side-to-side differences were also calculated. In this sense, when side-to-side difference > 6º was found, players were categorised as showing bilateral asymmetries whereas scores ≤ 6º were accepted as normal (non-bilateral asymmetries).[93]

5.3.3.7. Core stability

The unstable sitting protocol described by Barbado, Lopez-Valenciano, Juan-Recio, Montero-Carretero, van Dieen, & Vera-Garcia[260] was used to assess participants' ability to control trunk posture and motion while sitting. Briefly, after a familiarization/practice period (2 min), participants performed different static and dynamic tasks while sitting on an unstable seat:

- One static stability task without visual feedback (test 1) and another with visual feedback (test 2). In test 1, participants were asked to sit still in their preferred seated position on the unstable seat, while in test 2 participants were requested to adjust their centre of pressure position (CoP) to a target point located in the centre of a screen placed in front of them.
- Three dynamic stability tasks with visual feedback, in which participants were asked to track the target point, which moved along three possible trajectories (anterior-posterior, medial-lateral and circular).

All tasks were performed twice. The duration of each trial was 70 s and the rest period between trials was 1 minute. Participants performed each trial with arms crossed over the chest. All participants were able to maintain the sitting position without grasping a support rail.

The mean radial error was used as a global measure to quantify the trunk/core performance during the trials. This variable was calculated as the mean of vector distance magnitude of the CoP from the target point trials (trials with visual feedback) or from the participant's own mean CoP position (trials without visual feedback).[261]

5.3.3.8. Isokinetic knee flexion and extension strength

A Biodex System-4 isokinetic dynamometer (Biodex Corp., Shirley, NY, USA) and its respective manufacture software were used to determine isokinetic concentric and eccentric torques during knee extension and flexion actions in both legs following the methodology described by Ayala et al.[262]

Two isokinetic gravity-corrected variables were extracted for each movement (flexion and extension), muscle action (concentric, eccentric) and velocity (60, 180 and 240º/s for concentric actions and 30, 60 and 180º/s for eccentric actions): peak torque (PT) and joint angle of peak torque (APT). In each of the three trials at each velocity, the PT and APT were reported as the single highest torque output and corresponding joint angle. For each isokinetic variable, the average of the 3 sets at each velocity was used for subsequent statistical analysis. When a variation > 5% was found in the PT and APT values between the three trials, the mean of the two most closely related torque values were used for the subsequent statistical analyses.

Reciprocal (conventional and functional) knee flexion to knee extension ratios as well as bilateral knee flexion and extension ratios were also calculated using peak torque values extracted for each velocity. Thus, the conventional knee flexion to knee extension ratios were calculated as the ratio between the PTs produced concentrically by knee flexor and knee extensor muscles during the isokinetic tests. Functional knee flexion to knee extension ratios were calculated as the ratio between the PTs produced eccentrically by the knee flexor muscles and concentrically by the knee extensor muscles. Bilateral knee flexion and extension ratios were calculated dividing the PT value of the dominant leg by the PT value of the non-dominant leg. Finally, the functional knee flexion to knee extension ratio proposed by Croisier et al.,[118] was also calculated as the ratio between the PTs produced eccentrically by the knee flexor at 30º/s and concentrically by the knee extensor muscles at 240º/s.

5.3.4. Injury surveillance

Following the recommendations made by the international injury consensus group,[19] a MUS_{INJ} was defined as acute pain in the muscle location that occurred during training or match and resulted in the immediate termination of play and inability to participate in the next training session or match. These injuries were confirmed through a clinical examination (identifying pain on palpation, pain with isometric contraction, and pain with muscle lengthening) by team doctors. Players were considered injured until the club medical staff (medical doctor or physiotherapist) allowed full participation in training and availability for match selection. Only hamstrings, quadriceps, triceps surae and adductor muscles injuries were considered in this study.

The club medical staff of each club recorded MUS_{INJ} on an injury form that was sent to the study group each month. For all MUS_{INJ} that satisfied the inclusion criteria, team medical staffs provided the following details to investigators: muscle (hamstrings, quadriceps, triceps surae and adductors), leg injured (dominant/non-dominant), injury severity based on lay off time from soccer or handball (slight/minimal [0-3 days], mild [4-7 days], moderate [8-28 days], and severe [> 28 days]), date of injury, moment (training or match), whether it was a recurrence (defined as an MUS_{INJ} that occurred in the same leg and during the same season as the initial injury), and total time taken to resume full training and match. At the conclusion of the 9 month follow up period, all data from the individual clubs were collated into a central database, and discrepancies were identified and followed up at the different clubs to be resolved. Some discrepancies among medical staff teams were found to diagnose minimal MUS_{INJ} and to record their total time lost. To resolve these inconsistencies in the injury surveillance process (risk of misclassification of the players), only MUS_{INJ} showing a time lost > 4 days (minor to severe) were selected for the subsequent statistical analysis.

5.3.5. Statistical analysis

The statistical analysis framework carried out in this study for analysing and comparing the behaviours of several machine learning techniques with the aim of finding the best model for predicting MUS_{INJ} in professional soccer and handball players was based on a supervised learning perspective. From a statistical standpoint, the problem can be stated as follow: given a set of features F (in our case risk factors) and a target (discrete) variable (in our case MUS_{INJ} [yes or no]), named class C, we wanted to estimate/learn a map function $M{:}F \rightarrow C$. Thus, the statistical analysis comprised two stages:

1. Data pre-processing. At this stage, the data set was prepared to apply the data mining techniques. To optimise this aspect, pre-processing methods such as data cleaning and data discretization were applied.
2. Data processing. At this stage, the taxonomy suggested by Galar, Fernandez, Barrenechea, Bustince, & Herrera[113] to address learning with imbalanced data sets was applied. In particular, a study on the performance of some proposals for pre-processing, cost-sensitive learning and ensemble-based methods was carried out. In addition, the approach proposed by Elkarami, Alkhateeb, & Rueda[263] for imbalanced data set and based on the combination of a cost-sensitive classifier with class-balanced ensembles was also studied. Four classic decision tree algorithms were used as base classifiers in each method.

5.3.5.1. Data pre-processing

Data pre-processing is a crucial task, due to the quality and reliability of available information, which directly affects the results obtained. Thus, some specific pre-processing tasks were applied to prepare the data set so that the classification task could be carried out appropriately.

Firstly, we deleted those players who did not complete all the neuromuscular tests for any reason (six soccer players) from the data set. In addition, four soccer players were also deleted because they left their respective teams before the follow up procedure was completed. Secondly, we proceed to study the presence of outliers. In particular, we carried out an examination of the full data set using boxplots and the detected outliers were removed. The third step consisted of looking for missing data. To address this issue, frequency tables and diagrams were built. Thus, missing data were replaced by the mean value of the corresponding variable of the specific sport modality (soccer or handball) of the players. For example, if a soccer player did not report his weight for any reason, then the average value of his counterpart soccer players was inputted. It should be pointed out that none of the variables reported a percentage of missing data and outliers higher than 3%. The SPSS 21.0 Statistical software was used to carry out this data cleaning process.

After having applied the above-mentioned data cleaning methods, we had to deal with an imbalance data set (showing an imbalance ratio of 0.34) comprised of 88 soccer and 34 handball players (instances) and 151 potential risk factors (features).

The final step comprised the discretization of the continuous features as it has shown to be an effective measure to improve the performance of some classifiers.[264] Thus, continuous features were discretized according to the reference values previously reported to consider an player as being more prone to suffer an injury. In most features, the discretization reduced their dimensionality to three labels. In case no cut-off scores for detecting players at high risk of injury had been previously reported (e.g. stature, body weight, some isokinetic strength features), the unsupervised discretization algorithm available in the well-known Weka (Waikato Environment for Knowledge Analysis) data mining software was applied using the equal frequency binning approach (four cut point intervals). We selected four intervals in order to reflect taxonomy of low, low-moderate, moderate-high and high scores that might make the final model more comprehensible. For the discretization of the psychological features (Appendix 5.2) and the isokinetic APT features we used two and three intervals or labels respectively setting up according to the authors´ extensive experience due to the fact that their range of possible scores were limited (e.g. from 0 to 5). In this sense, lower extremity ROM features (Appendix 5.3) as well as both reciprocal knee flexion to knee extension ratios and bilateral knee flexion and extension ratios (Appendix 5.4) were discretised according to the previously suggested cut-off scores whereas dynamic postural control (Appendix 5.5), isometric hip abduction and adduction strength (Appendix 5.6), core stability (Appendix 5.7) and isokinetic peak torque (Appendix 5.4) features were discretized using the Weka unsupervised discretization algorithm.

5.3.5.2. Data processing

Although in data mining and machine learning a wide range of paradigms have been used to tackle classification problem, only those that have designed to deal with imbalance datasets were used. These paradigms might be categorized into three groups:[113,114]

a) External approaches that pre-process the data in order to reduce the effect of their class imbalance by resampling the data space.

b) Internal approaches that create new algorithms or modify existing ones to take the class imbalance problem into consideration (ensembles).

c) Cost-sensitive learning solutions incorporating both the data (external) and algorithmic level (internal) approaches assume higher misclassification costs for samples in the minority class and seek to minimize the high cost errors.

The taxonomy for external (oversampling), internal (ensembles) and cost-sensitive methods for learning with imbalanced data sets proposed by Galar et al.,[113] and López, Fernández, García, Palade, & Herrera[114] was used to address the aim of this study. This taxonomy was implemented with the approach recently proposed by Elkarami et al.,[263] due to the promising results showed to handle imbalanced data sets.

To achieve founded conclusions, four decision tree algorithms were selected to be used in the pre-processing, ensemble and cost sensitive learning methodologies: C4.5,[265] which is an algorithm for generating a pruned or unpruned decision tree; SimpleCart,[266] which implements minimal cost-complexity pruning; ADTree,[267] which is an alternating decision tree; and RandomTree,[268] which considers K randomly chosen attributes at each node of the tree.

In this sense, a decision tree is a set of conditions organized in a hierarchical structure. An instance is classified by following the path of satisfied conditions from the root of the tree until a leaf is reached, which will correspond with a class label.

For the sake of brevity and the lack of space, we have not written here the code of the algorithms used in this study. Instead, we have only specified the names and refer the reader to their original sources. Furthermore, all the classification algorithms used are available in Weka data mining software.

The use of balancing (e.g. oversampling and undersampling) techniques in the training data subset prior to running decision tree algorithms has been shown to be an effective measure to increase the performance of the latter when an imbalance data set is presented.[114] Although there are several data balancing or rebalancing algorithms, we used three of the most popular methodologies which are the synthetic minority oversampling technique (SMOTE), random oversampling (ROS) and random undersampling (RUS). In brief, its main idea behind SMOTE is to create new minority class examples by interpolating several minority class instances that lie together for oversampling the training set. With these techniques, the minority class is over-sampled by taking each minority class sample and introducing synthetic examples along the line segments joining any/all of the k samples belonging to the minority class, nearest to the sample *i*. Regarding ROS, it duplicates some random minority instances until the total amount of minority instances reaches the percentage given and RUS, contrarily, removes some random majority samples. In our case, a level of balance in the training data near to the 40:60 was tried to obtain. Additionally, the interpolations that are computed to generate new synthetic data were made considering the k-5-nearest neighbours of minority class instances using the Euclidean distance.

Regarding ensemble learning algorithms, classic ensembles such as Bagging, AdaBoost and AdaBoot.M1 were included in this study. Further, the algorithm families designed to deal with skewed class distributions in data sets were also included: Boosting-based and Bagging-based. The Boosting-based ensembles that were considered in the current study were SMOTEBoost and RUSBoost. About bagging-based ensembles, it was included from the OverBagging group, OverBagging (which uses random oversampling), UnderBagging (which uses random undersampling) and SMOTEBagging.

Concerning the cost-sensitive learning algorithms, two different approaches were used, namely metacost and the cost sensitive classifier. Instead, we have only specified the names and refer the reader for further information to Galar et al.,[113] and López et al.[114] Regarding the number of internal classifiers used within each approach, all ensembles employed 10 base classifiers by default.

Finally, the behaviour of some specific combination of class-balanced ensembles with cost-sensitive base classifiers was also studied. The final cox matrix set up was based on the best performance reported after testing all the possibilities. Appendix 5.8 summarizes the list of algorithms grouped by families and also shows the abbreviations that have been used along the experimental framework and a short description of them.

To evaluate the performance of the decision tree algorithms, the worldwide-accepted fivefold stratified cross validation (SCV) technique was used.[269] That is, we split the dataset into five stratified folds maintaining the class distribution, each one containing 20% of the patterns of the dataset. For each fold, the algorithm was trained with the examples contained in the remaining folds and then tested with the current fold. This value is set up with the aim of having enough positive class instances in the different folds, hence avoiding additional problems in the data distribution. A wide range of classification performance measures can be obtained from the SCV technique. A well-known approach to unify these measures and to produce an evaluation criterion is to use the Receiver Operating Characteristic (ROC) curve. In particular, the area under the ROC curve (AUC) corresponds to the probability of correctly identifying which one of the two stimuli is noise and which one is signal plus noise.[114] Thus, the AUC was used as a single measure of a classifier's performance for evaluating which model is better on average and was interpreted as high (0.90- 1.00), moderate (0.70-0.90), low (0.70-0.50), and fail (>0.50).[270] Furthermore, two extra measures from the confusion matrix were also used as evaluation criteria: a) true positive rate (TPrate): TPrate = $\frac{TP}{TP + FN}$ also called sensitivity or recall, is the proportion of actual positives which are predicted to be positive; and b) true negative rate (TNrate): TNrate = $\frac{TN}{TN + FP}$ or specificity, that is the proportion of actual negatives which are predicted to be negative.

5.4. Results

5.4.1. Muscle injuries epidemiology

There were 32 $\mathrm{MUS_{INJ}}$ over the follow up period, 21 (65.6%) of which corresponded to the hamstrings, four to the adductors (12.5%), four to the triceps surae (12.5%) and three to the quadriceps (9.3%). Injury distribution between the legs was 53.3% dominant leg and 46.7% non-dominant leg. A total of 13 injures occurred during training and 19 during match. In term of severity, most injures were categorized as moderate (n = 23) while only 9 cases were considered minor and no severe injuries were recorded. Three players were injured twice during the observation period, and only their first injury was used as the index injury in the analyses. Consequently, 29 $\mathrm{MUS_{INJ}}$ were finally used to develop the predictive models.

5.4.2. Predictive model for lower extremity muscle injuries

Tables 5.1-5.3 show the average AUC, TPrate and TNrate results for all resampling, ensemble and cost-sensitive learning methods separately for each decision tree base classifier.

Table 5.1. Average area under the receiver operating characteristic curve, true positive rate and true negative rate results for all the decision tree methodologies in isolation and after having been applied in them the resampling techniques selected.

Technique	**AUC**	TPrate	TNrate
Base classifiers			
J48	0.422	17.2	79.1
SCart	0.462	3.4	94.5
ADTree	**0.623**	**20.7**	**87.9**
RTree	0.609	51.7	65.9
Oversampling techniques			
SMT			
J48	0.452	31	78
SCart	0.489	34.5	71.4
ADTree	**0.608**	**31**	**76.9**
RTree	0.522	34.5	71.4
ROS			
J48	0.575	44	72.5
SCart	0.618	48.3	73.6
ADTree	**0.709**	**48.3**	**84.6**
RTree	0.711	55.2	82.4
Undersampling techniques			
RUS			
J48	0.607	55.2	62.4
SCart	0.574	13.8	93.4
ADTree	**0.662**	**62.1**	**70.3**
RTree	0.559	48.3	61.5

Abbreviations can be found in appendix 5.8. The method that obtained the best performing result within each method is highlighted in bold.

The ADTree base classifier showed the best performance in most of the methods analysed. In fact, the final model was built using the SMOTEBagging ensemble method with the ADTree as base classifier using reweighted training instance (cost-sensitive).

Table 5.2. Average area under the receiver operating characteristic curve, true positive rate and true negative rate results for the ensembles techniques.

Technique	**AUC**	TPrate	TNrate
Classic Ensembles			
ADB1			
J48	0.579	13.8	90.1
SCart	0.605	37.9	83.5
ADTree	**0.692**	**24.1**	**93.4**
RTree	0.594	10.3	98.9
M1			
J48	0.560	0	91.2
SCart	0.550	20.7	84.6
ADTree	**0.703**	**27.6**	**90.1**
RTree	0.517	20.7	85.7
BAG			
J48	0.544	6.9	93.4
SCart	0.669	3.4	97.8
ADTree	**0.722**	**10.3**	**98.9**
RTree	0.663	24.1	91.2
Boosting-based Ensembles			
SBO			
J48	0.494	24.1	76.9
SCart	**0.692**	**41.4**	**85.7**
ADTree	0.650	27.6	85.7
RTree	-	-	-
RUSB			
J48	0.610	37.9	75.8
SCart	0.649	51.7	78
ADTree	0.698	31	92
RTree	**0.717**	**48.3**	**84.6**
Bagging-based Ensembles			
OB			
J48	0.583	13.8	92.3
SCart	0.716	13.8	93.4
ADTree	**0.759**	**10.3**	**96.7**
RTree	0.633	13.8	89.0
UB			
J48	0.670	27.6	84.6
SCart	**0.708**	**31**	**87.9**
ADTree	0.624	41.4	73.6
RTree	0.570	27.6	82.4
SBAG			
J48	0.562	13.8	96.7
SCart	0.642	10.3	96.7
ADTree	**0.728**	**20.7**	**96.7**
RTree	0.547	24.1	93.4

Abbreviations can be found in appendix 5.8. The method that obtained the best performing result within each method is highlighted in bold.

Table 5.3. Average area under the receiver operating characteristic curve, true positive rate and true negative rate results for the cost-sensitive learning and class-balanced ensembles with cost-sensitive classifier techniques.

Technique	**AUC**	TPrate	TNrate
Cost-sensitive classification			
MetaCost			
J48	0.473	41.4	61.5
SCart	0.579	17.2	90.1
ADTree	**0.662**	**75.9**	**40.7**
RTree	0.561	48.3	63.7
CS-Classifier			
J48	0.526	51.7	57.1
SCart	0.543	44.0	52.7
ADTree	**0.642**	**51.7**	**70.3**
RTree	0.535	44.0	60.4
Class-balanced ensembles with a cost-sensitive classifier			
CS-SBAG			
J48	0.529	51.7	51.6
SCart	0.610	65.5	54.9
ADTree	**0.747**	**65.5**	**79.1**
RTree	0.541	6.9	86.8
CS-OBAG			
J48	0.514	41.4	72.5
SCart	0.606	55.2	63.7
ADTree	**0.742**	**62.1**	**71.4**
RTree	0.548	13.8	96.7
CS-UBAG			
J48	0.553	41.4	67
SCart	0.649	51.7	69.2
ADTree	**0.742**	**58.6**	**68.1**
RTree	0.627	37.9	82.4

Abbreviations can be found in appendix 5.8. The method that obtained the best performing result within each method is highlighted in bold. The model considered as the best for predicting muscle injuries is highlighted in grey.

Therefore, the final model selected to predict lower extremity MUS_{INJ} in professional soccer and handball players is comprised by 10 different cost sensitive classifiers (ADTrees) (Figures 5.2-5.11) and 52 features (Appendix 5.9). The cost matrix for cost-sensitive classifier was set to

$$C\left\{\frac{0}{2}\middle|\frac{14}{0}\right\}$$

where a false negative had a cost of 14 and a false positive had a cost of 2. In our case, the false prediction of a non-injured player was penalized seven times more with respect to the contrary error.

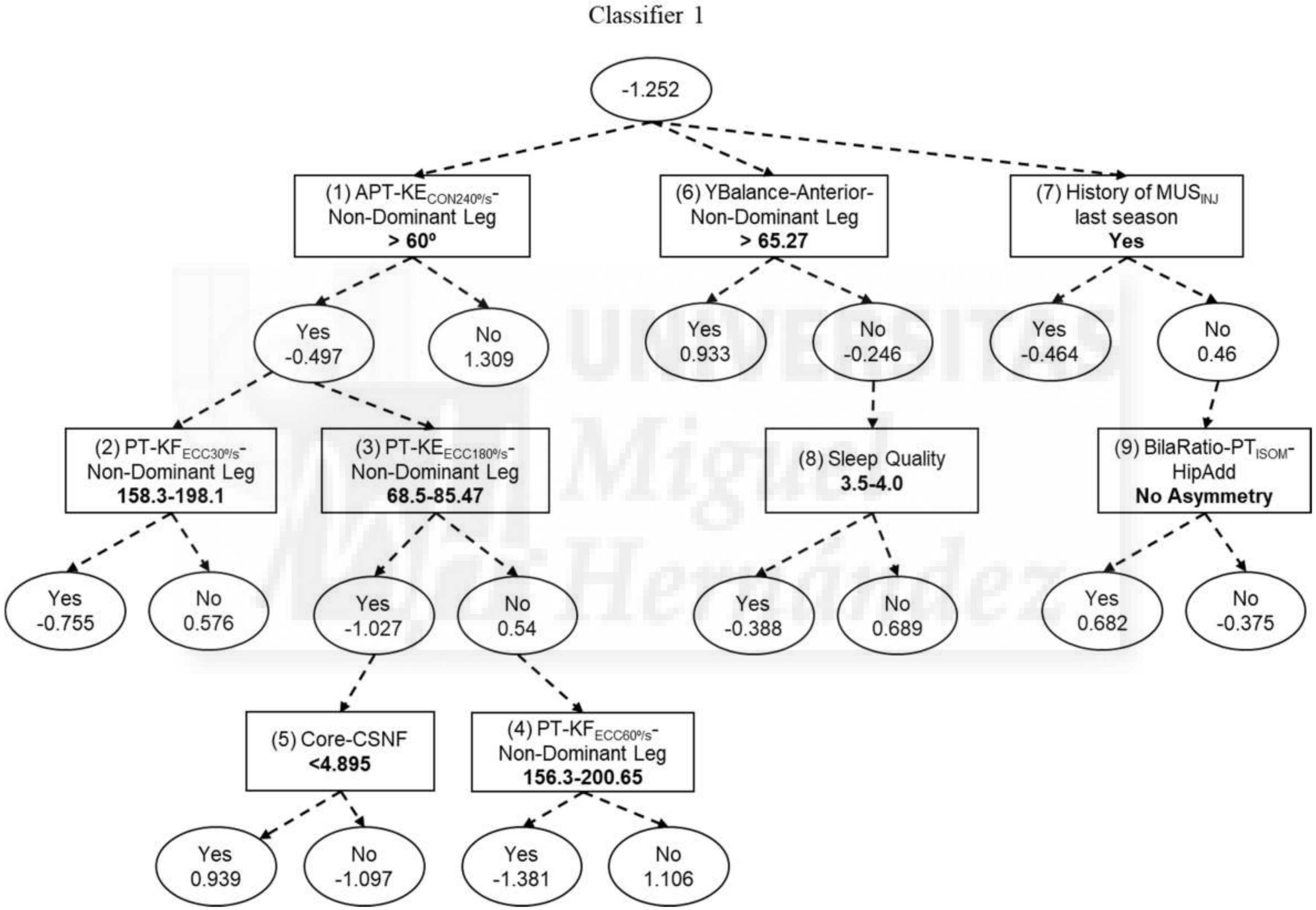


Figure 5.2. Muscle injuries predictive model, classifier 1.

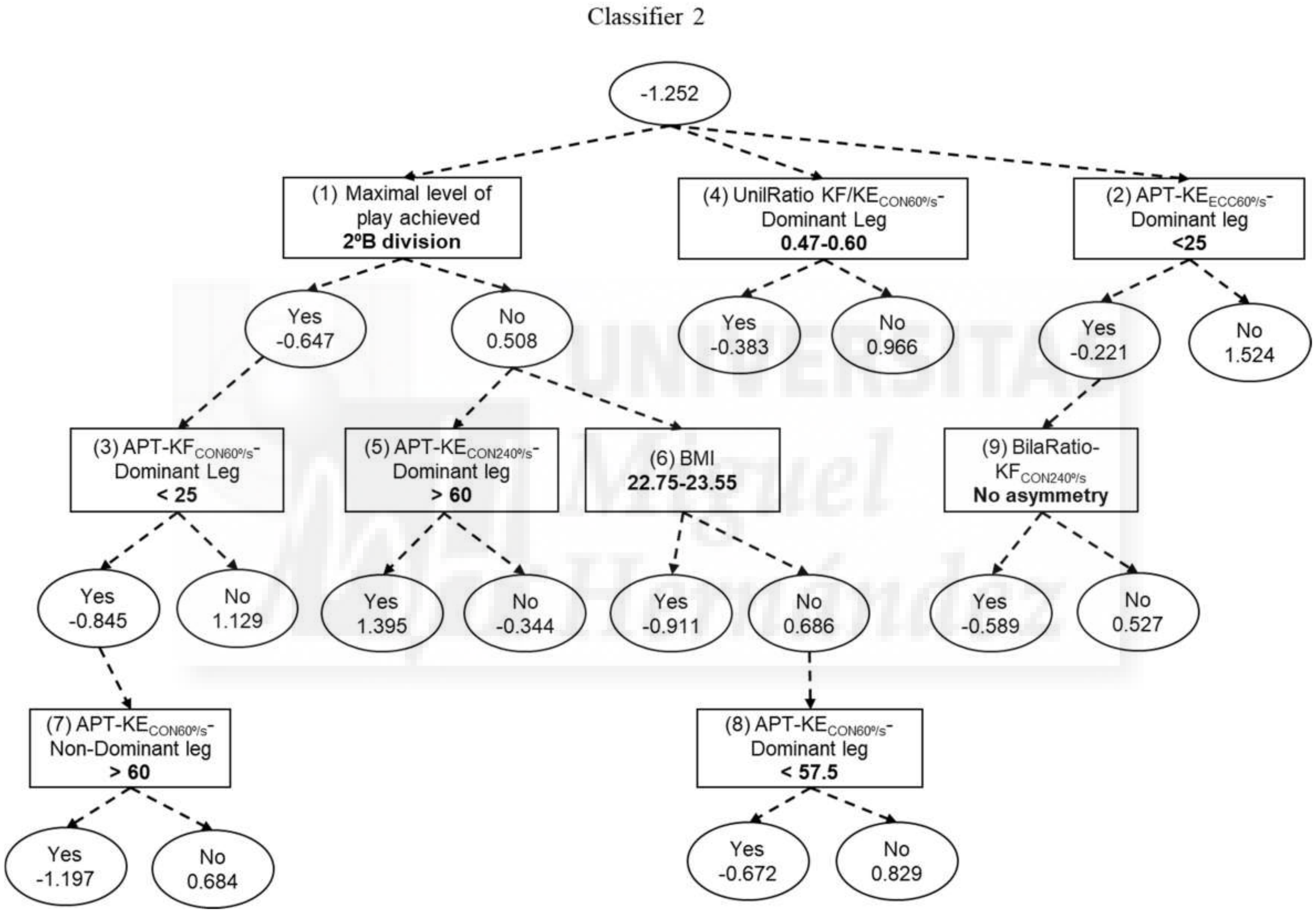


Figure 5.3. Muscle injuries predictive model, classifier 2.

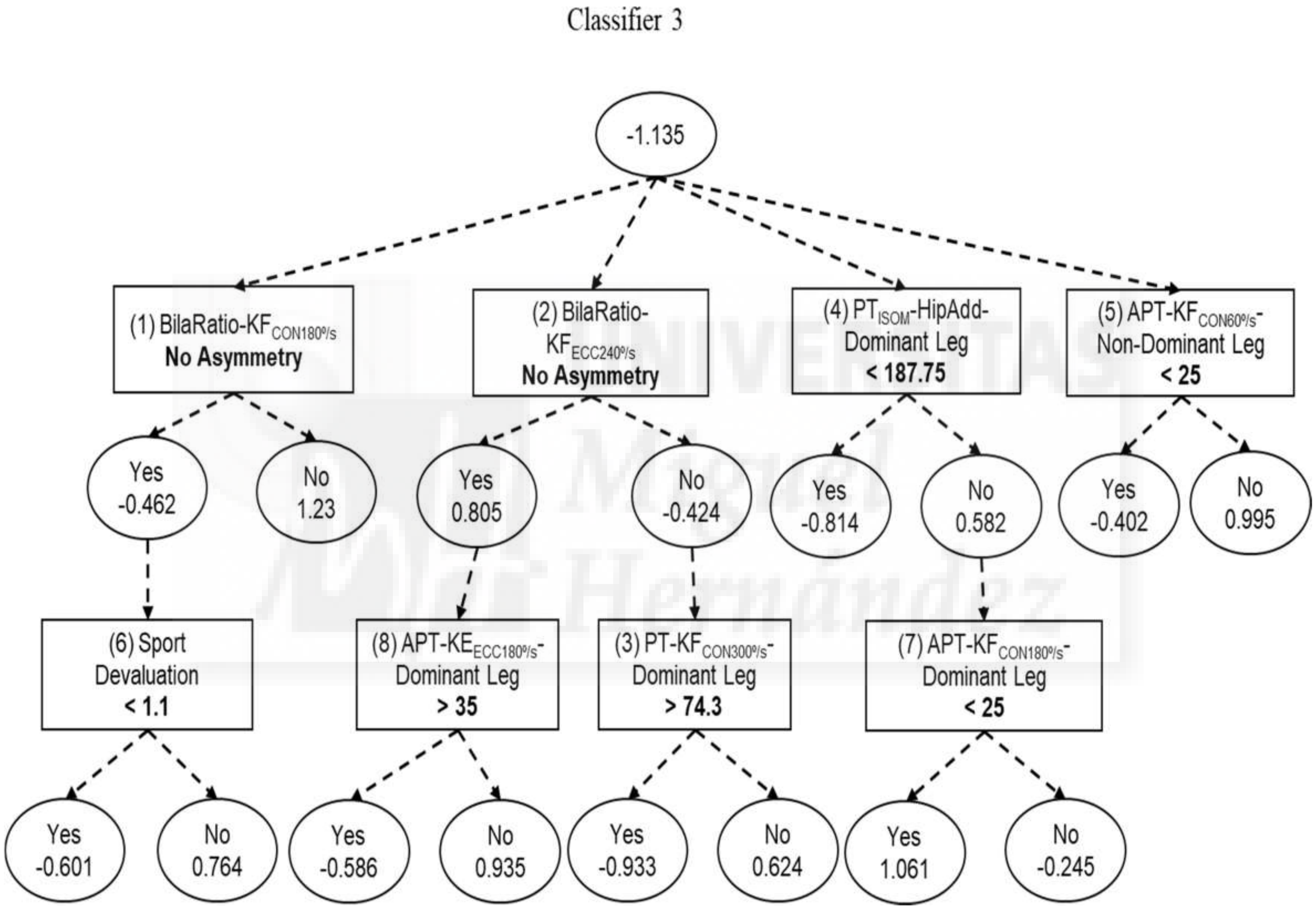


Figure 5.4. Muscle injuries predictive model, classifier 3.

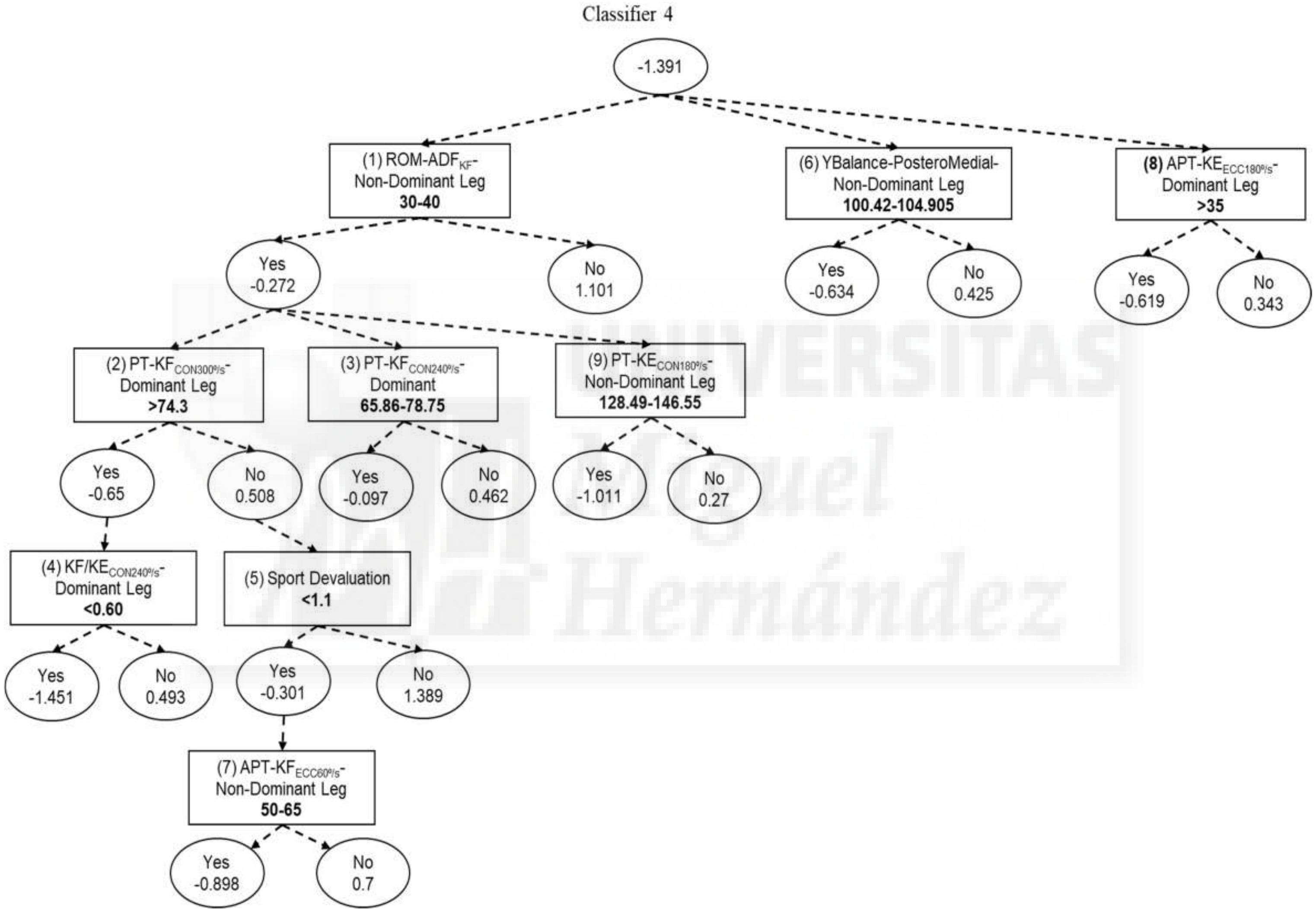


Figure 5.5. Muscle injuries predictive model, classifier 4.

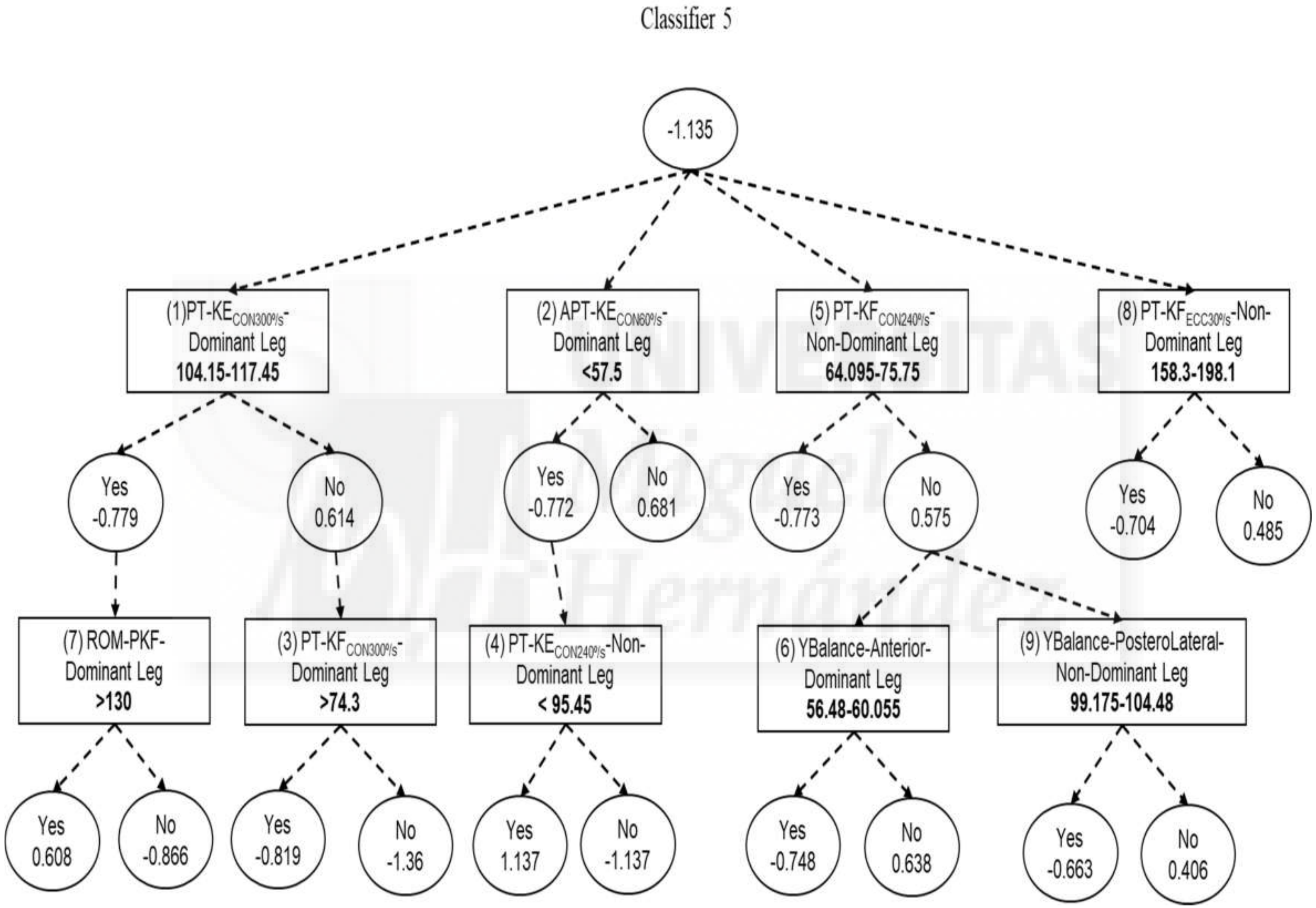


Figure 5.6. Muscle injuries predictive model, classifier 5.

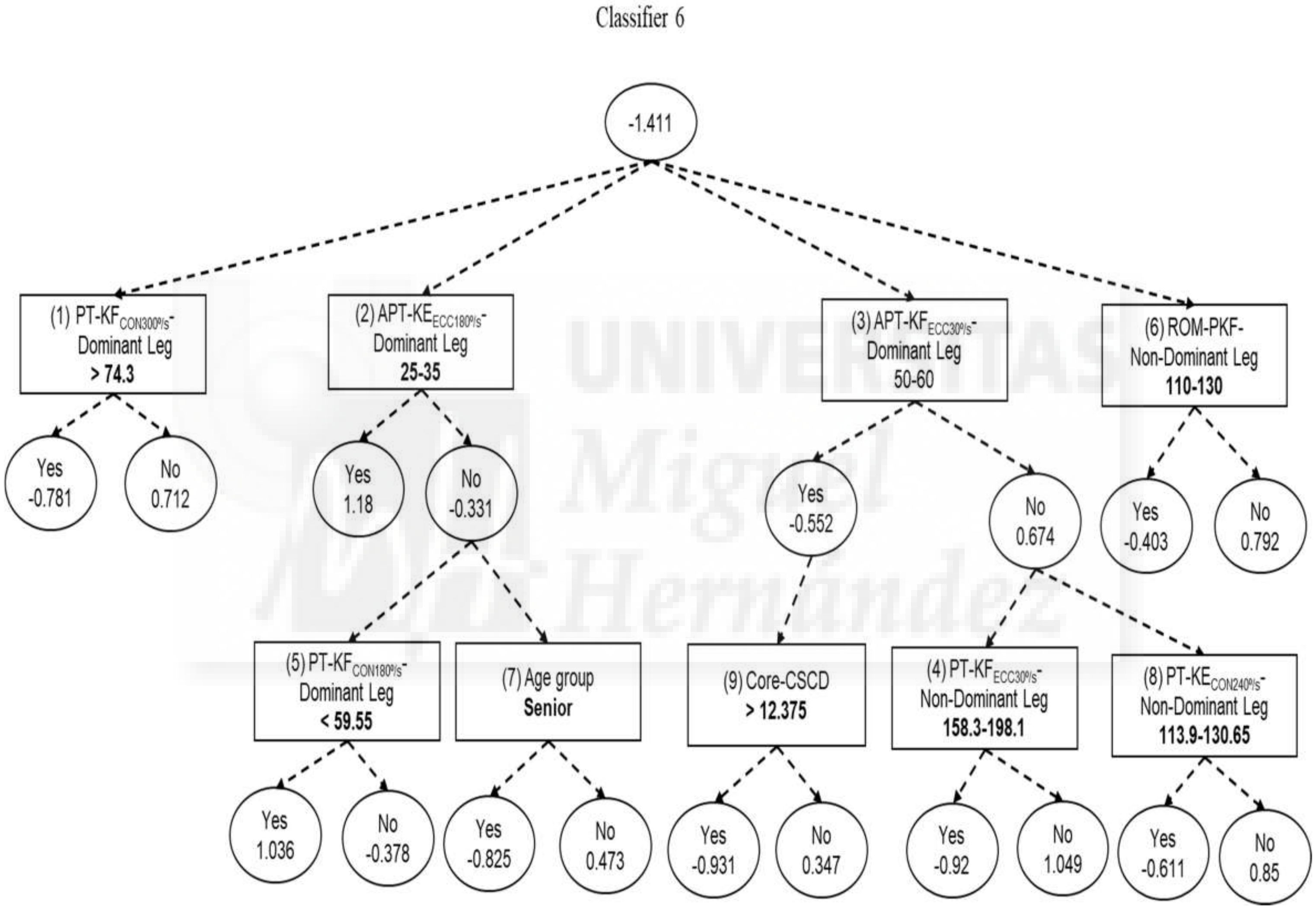


Figure 5.7. Muscle injuries predictive model, classifier 6.

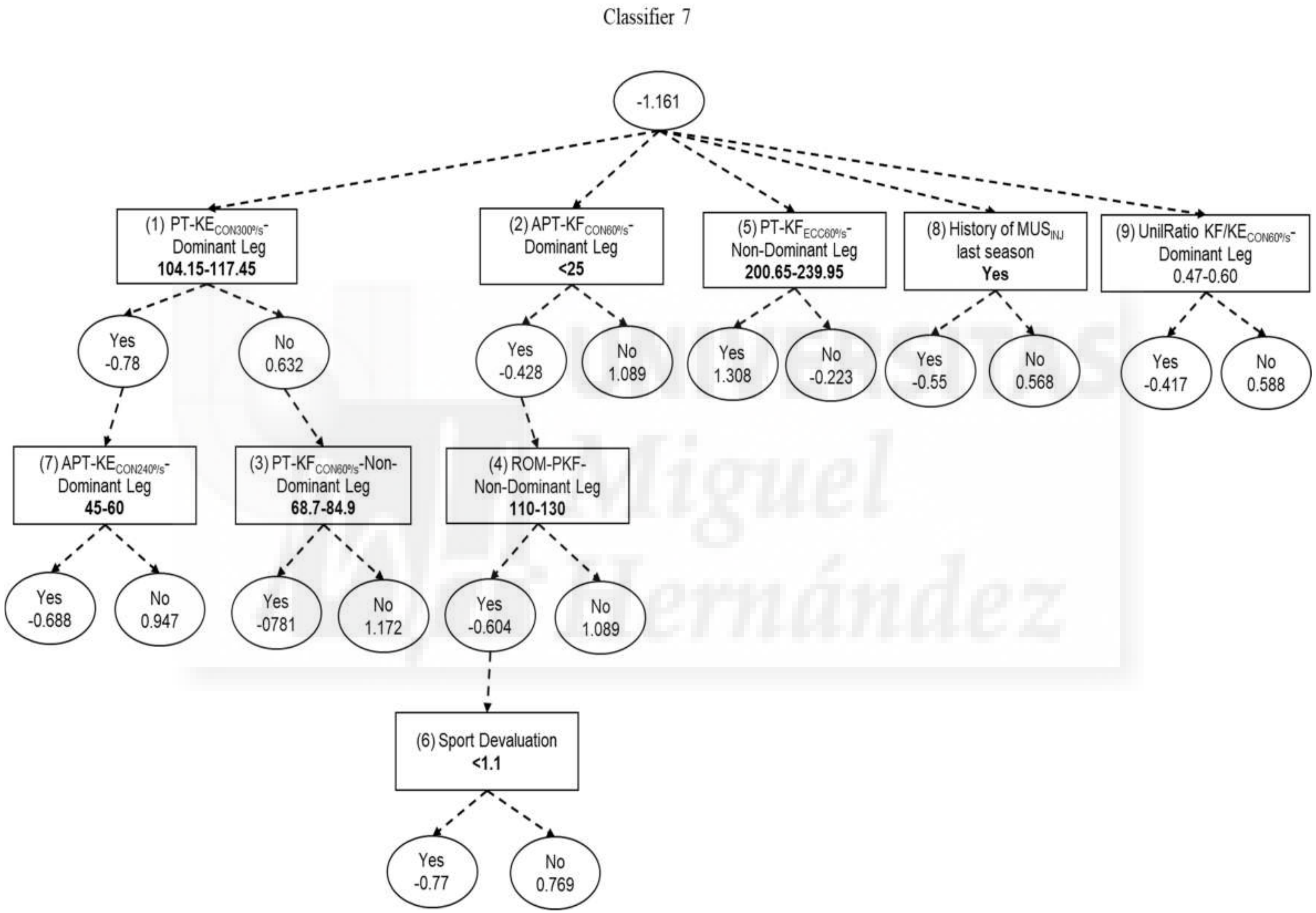


Figure 5.8. Muscle injuries predictive model, classifier 7.

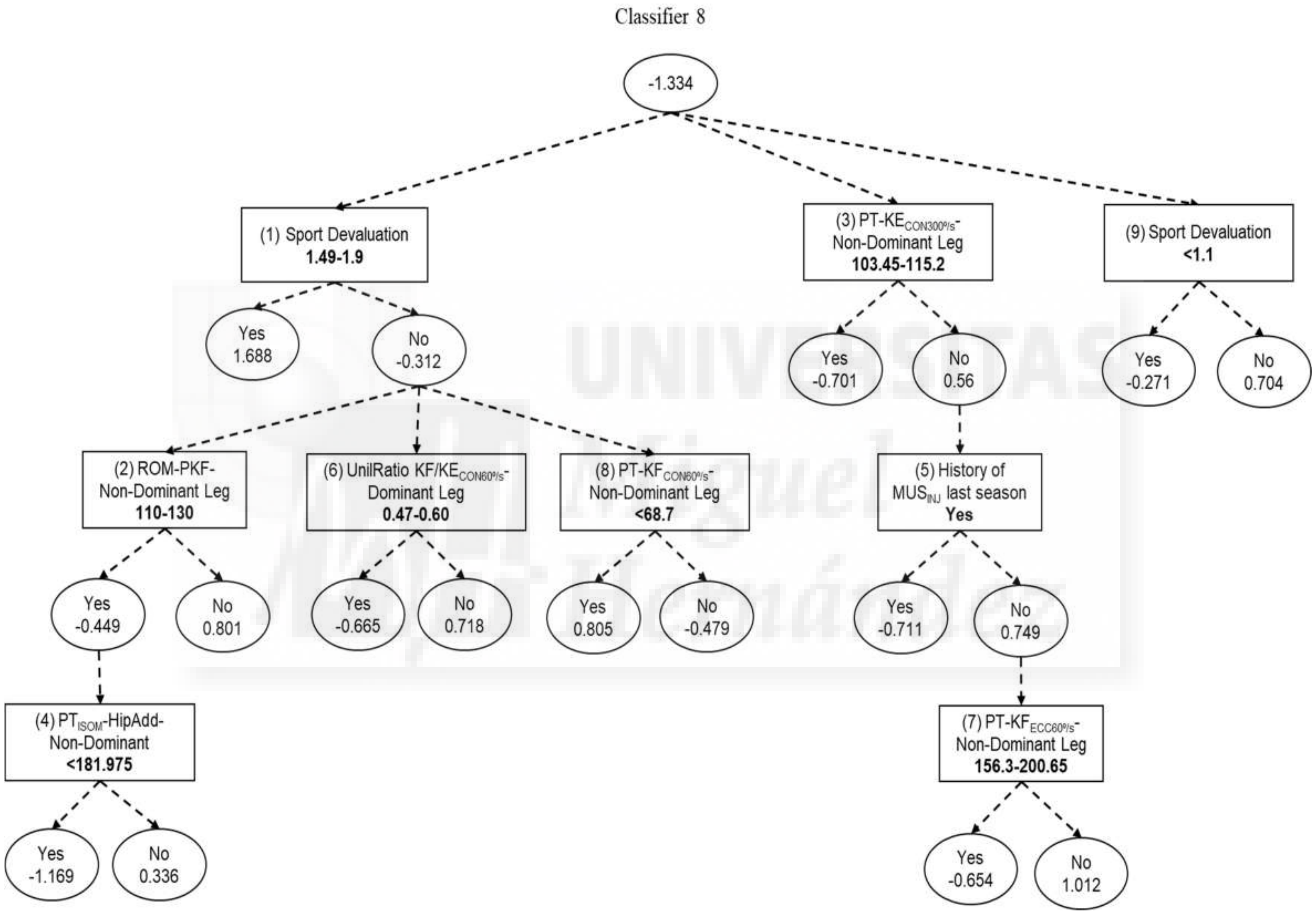


Figure 5.9. Muscle injuries predictive model, classifier 8.

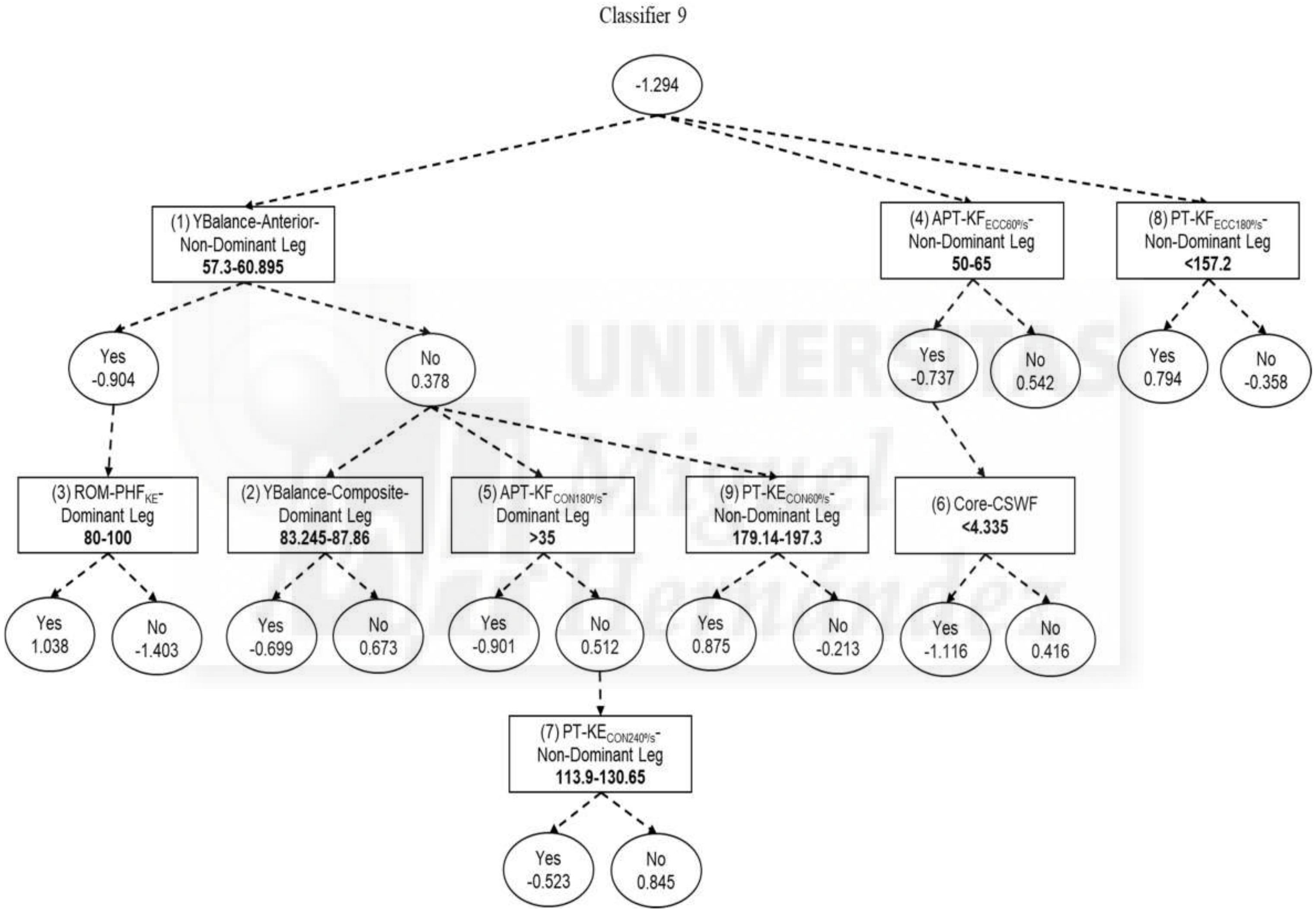


Figure 5.10. Muscle injuries predictive model, classifier 9.

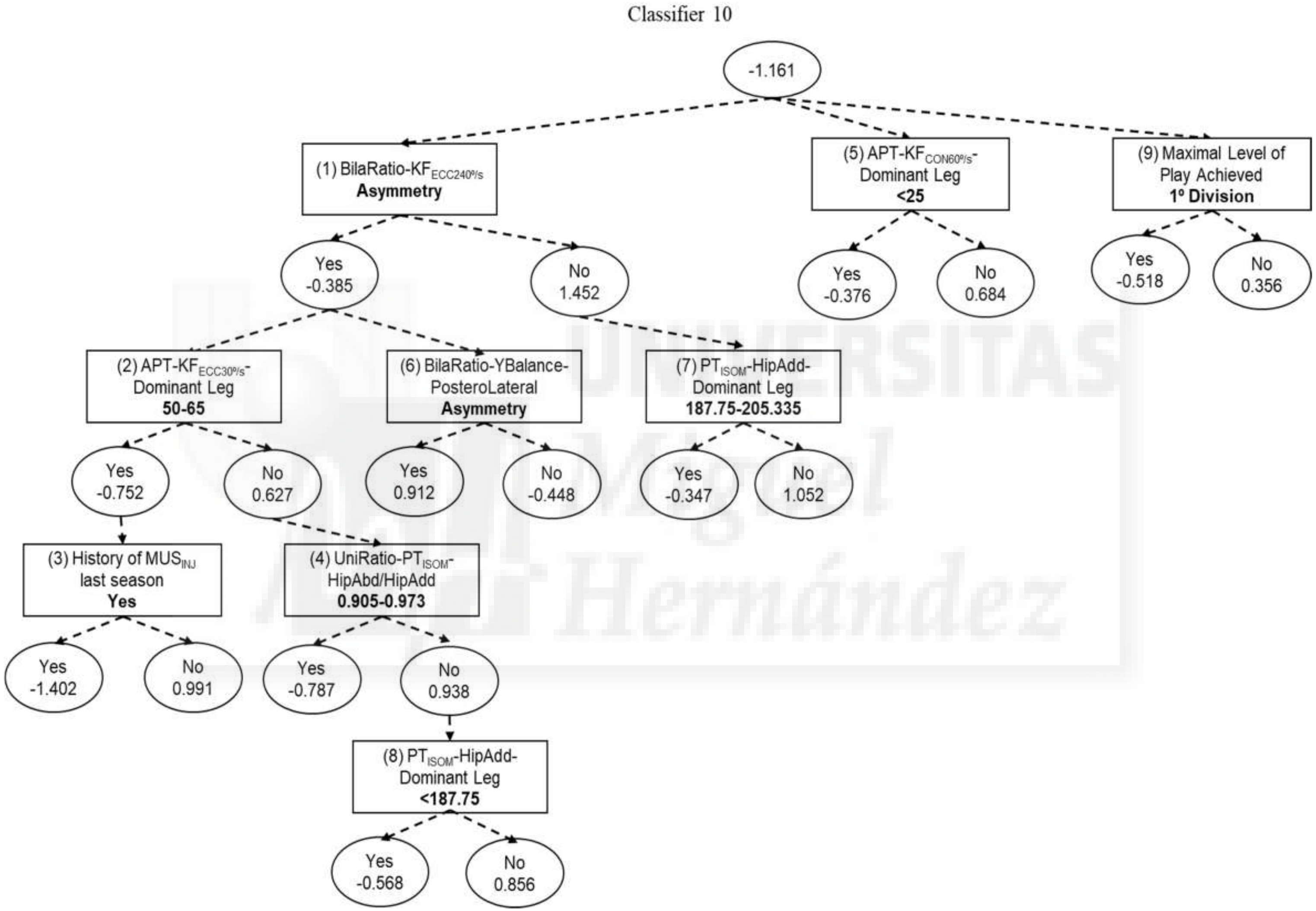


Figure 5.11. Muscle injuries predictive model, classifier 10.

The confusion matrix and the main cross validation results of the final model are shown in tables 5.4 and 5.5. In terms of practical applications, each classifier has a vote (yes or no), and the final decision regarding whether or not a player might suffer an injury will be based on the combination of the votes of each individual classifier to each class (yes or no).

Table 5.4. Confusion matrix.

A	B	Classified as
19	10	A = Injured
19	72	B = Not Injured

Table 5.5. Cross validation results for the final prediction model.

Correctly classified instances	91 (75.8%)
Incorrectly Classified Instances	29 (24.1%)
Kappa statistic	0.401
Mean absolute error	0.405
AUC	0.747

AUC: area under the receiver operating characteristic curve.

5.5. Discussion

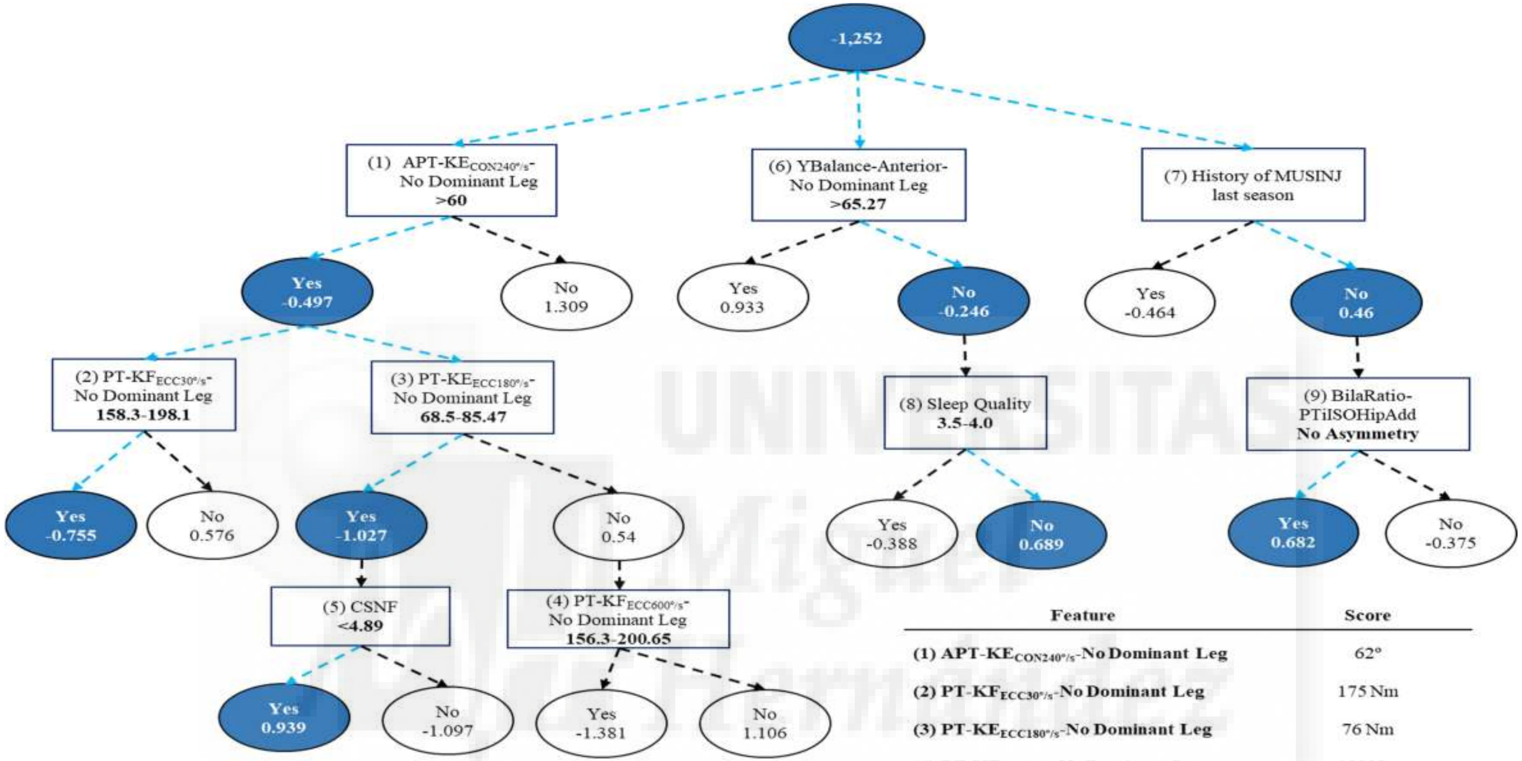
The main purpose of this study was to develop an injury risk factor-based model that would identify professional soccer and handball players at high risk of MUS_{INJ} by using learning methods coming from machine learning and data mining environments. With this aim in mind, a large number of personal, psychological and neuromuscular risk factors were assessed during the preseason training periods and the MUS_{INJ} accounted within the following 9 months were also recorded. Thus, and after having run and compared the performance of several pre-processing, cost-sensitive learning and ensemble techniques to correctly classify players at high or low risk of MUS_{INJ}, the model generated by the SmooteBoost technique with a cost-sensitive ADTree as base classifier reported the best evaluation criteria (AUC score = 0.747; TPrate = 65.9; TNrate = 79.1).

5.5.1. Functioning of the predictive model to identify players at high risk of muscle injuries

The ADTree algorithm has the advantage of producing models that are easily represented as a tree with a limited number of nodes (less than 10 in our case). This property is achieved by constructing a tree that is a conjunction of rules which all contribute real-valued evidence toward a given instance being classified as either true (injured) or false (not injured). Unlike traditional tree models the classification of instances by ADTree is thus not determined by a single path traversed in the tree, but rather by the additive score of a collection of paths. The ADTree is graphically represented with two types of nodes: Elliptical *prediction nodes* and rectangular *splitter nodes* (Figures 5.2-5.11). Each splitter node is associated with a value indicating the rule condition: If the feature represented by the node satisfied the condition for a given instance, the prediction path will go through the left child node, otherwise the path will go through the right child node. The final classification score produced by the tree is found by summing the values from all the prediction nodes reached by the instance, with the root node being the precondition of the classifier. If the summed score is greater than zero, the instance is classified as false (not injured).

To better explain how coaches and sport practitioners should use the model to predict MUS_{INJ}, we are going to explain the first classifier or ADTree using the fictional data displayed in figure 5.12. In addition, figure 5.12 represents in blue the paths followed by the selected instance or example.

In this classifier, we start with a baseline score of -1.252. The tree presents three father nodes placed up to the tree: APT_{ISOK}-$\text{KE}_{\text{CON240º/s}}$-Non-dominant Leg, YBalance-Anterior-Non-dominant Leg and History of MUS_{INJ} last season. Each father node represents a pathway that must be addressed.



Feature	**Score**
(1) APT-KE$_{CON240°/s}$-No Dominant Leg	62°
(2) PT-KF$_{ECC30°/s}$-No Dominant Leg	175 Nm
(3) PT-KE$_{ECC180°/s}$-No Dominant Leg	76 Nm
(4) PT-KF$_{ECC60°/s}$-No Dominant Leg	193 Nm
(5) Core-CSNF	3.78
(6) YBalance-Anterior-No Dominant Leg	62 cm
(7) History of MUS$_{INJ}$ last season	No
(8) Sleep Quality	4.2
(9) BilaRatio-PT$_{ISOM}$-HipAdd	0.98 (No Asymmetry)

Figure 5.12. Graphical representation of the first classifier. Prediction nodes are represented by ellipses and splitter nodes by rectangles. Each splitter node is associated with a real valued number indicating the rule condition, meaning: If the feature represented by the node satisfies the condition value the prediction path will go through the left child node, otherwise the path will go through the right child node. The numbers before the feature names in the prediction nodes indicate the order in which the different base rules were discovered. This ordering can to some extent indicate the relative importance of the base rules. Abbreviations can be found in appendixes 5.3-5.7.

Then, and if we start by the father node numbered as 1, placed on the left and represented by the feature named APT_{ISOK}-$KE_{CON240º/s}$-Non-dominant Leg, we realise that our player satisfies the rule condition, this is, he presents a score > 60º (Yes). Consequently, we must sum -0.497 to the initial score. Then, we have two different pathways that must be addressed. Thus, we first address the pathway that goes toward the node that contains the feature named PT_{ISOK}-$KF_{ECC30º/s}$-Non-dominant Leg. Our player satisfies again the rule condition (Yes) because he shows a score ranged from 158.3 to 198.1. Therefore, we sum -0.755 to the baseline score. Until here, we have reached an accumulative score of -2.504 (-1.252 + [- 0.497] + [- 0.755]).

If we go back to the node number 1, and we follow the remaining pathway that goes toward the node number 3, we check that our player satisfies its rule condition, and then we add other -1.027 points to our scoreboard (-2.504 + [-1.027] = -3.531). As the path is not finished, we must continue through the Yes path and reach the last node, represented by the feature Core-CS_{NF}. Here, our player satisfies again the rule condition and we must sum 0.939 point to our accumulate scoreboard. It should be noticed that this time the score summed is positive and hence, our accumulative score would be reduced. Therefore, by completing this first pathway started in the node 1 we have reached a total score of -2.592. Once we have completed this first path we must proceed with the other two primary paths, but taking into account that we have an accumulative scoreboard of -2.592.

Thus, and after completing the second main pathway, we must sum -0.246 (YBalance-Anterior-Non-dominant Leg = No) and + 0.689 (Sleep Quality = No) points to our scoreboard. Finally, we also have to sum 0.46 and 0.682 points coming from the third main pathway. All in all, our player has reached a global score of -1.007. The higher the global score is (in positive or negative way), the more confidence we can show with the vote obtained

Consequently, this classifier votes "Yes" and thus considers our player at high risk of injury. The final classification will be based on the combination of the votes of each individual classifier to each class (yes or no). In the very unlikely (but possible) case where a player would end in a draw between votes of different sign (i.e. five votes for no and five votes for yes), coaches and sport practitioners should adopt a conservative attitude and consider the player at high risk of MUS_{INJ}. The rationale behind of this recommendation for the unlikely case of draw is based on the reported high incidence rate of muscle injuries in professional sports[41,251,252] and on the cost that a false negative diagnosis (low sensitivity) might have for team performance and player´s welfare as well as the economic cost for the club.[159,176]

5.5.2. Discussion of the predictive model results

As it has been stated before, the model generated is comprised by 10 classifiers that contain the most relevant features (n = 52) for predicting MUS_{INJ}. In addition, each feature presented in the model shows a binary rule condition (yes or no) based on a specific cut-off score. Therefore, we consider that the model meets the two requirements (i.e. identifying relevant risk factors and defining cut-off scores) established in the first step suggested by Bahr [7] to be considered as a valid screening methodology.

In this sense, the predictive model built considers the devaluation of the self-perceived benefits gained from sport involvement as being one of the main factors associated with an increased in the relative risk of MUS_{INJ} because it is presented in 5 of the 10 classifiers. This finding is in concordance with the results found by Cresswell, & Eklund,[271] who reported statistically significant correlations between sport-injuries and feelings of sport devaluation in a cohort of professional rugby players. Although the mechanisms behind the relationship between sport devaluation and injury have not been well defined yet, it might be possible that old professional players with a short term history of moderate to severe injuries would start questioning if the efforts made to achieve their current level of play deserve the benefits gained. These feelings of frustration might lead players to loss the concentration and reduce the intensity of their actions during both training and match play, and thus increasing the risk of MUS_{INJ}. Therefore, psychological therapies aimed at reducing player burn out could help to reduce the risk of MUS_{INJ} in professional soccer and handball players.

Another strong risk factor reported by the model (presented in four classifiers) for MUS_{INJ} is having a history of MUS_{INJ} last season. Previous injury has been also identified in some prospective studies as one of the primary risk factors for MUS_{INJ}.[20,23,57] A possible explanation for previous injury being such a consistent risk factor for re-injuries may be that the joints or muscles in question are not fully restored structurally and/or functionally.[57] Consequently, more studies are needed in order to: a) design effective rehabilitation programmes after injury; and b) develop adequate return-to-play guidelines. Furthermore, evidence-based MUS_{INJ} prevention programs should be applied since the beginning of each player´s sport career in order to avoid or postpone the first MUS_{INJ} as a high priority, in order to keep players from entering the vicious cycle of repeated injuries to the same muscle group.

Furthermore, the model built provides a main role to the isokinetic strength features measured through knee flexion and extension actions to predict future $MUSC_{INJ}$ (30 features up to 52). These results are not in agreement with the findings showed by van Dyk et al.,[125] who reported that the use of isokinetic testing to determine the association between strength differences and hamstring muscle injuries was not supported. A possible reason behind the discrepancy between the findings reported by van Dyk et al.,[125] and our results might be associated with the different statistical approach used. Thus, while van Dyk et al.,[125] carried out a clustered multiple logistic regression analysis to identify variables associated with the risk of hamstrings injuries, we used a more robust analysis that took into account the different distribution presents in the class feature. It should be highlighted that our model endows a special protagonist for predicting future $MUSC_{INJ}$ to the APT measured through concentric (quadriceps) and eccentric (hamstrings) knee extension movements, as they are presented in 4 and 5 different classifiers respectively. This circumstance might support the hypothesis derived from the findings reported by Brockett, Morgan, & Proske[117] so that where the players are able to achieve the PT might be more relevant than the net PT value in order to prevent $MUSC_{INJ}$.

On the other hand, another relevant isokinetic feature for our predictive model is the conventional knee flexion and extension ratio measured at 60º/s. Surprisingly, no functional knee flexion and extension ratio feature were included in the final models despite being more conceptually relevant for muscle injuries than the conventional ratios (mainly hamstrings injuries). Perhaps a potential reason for this circumstance might be based on the use of inappropriate cut-off scores. In this sense, we categorised the functional knee flexion and extension ratios using the cut-off scores reported in the literature. It is possible that these cut-off scores calculated using different isokinetic methodologies could not have been appropriate (very restrictive) for our model and hence, reduced its performance.

Although with less presence than the isokinetic features, the classifiers that compose the predictive model including features from all the testing methodologies used, which might support the multifactorial character of the MUS_{INJ} phenomenon. This characteristic of the model might support its congruence. Using the cross-validation process, we consider that the model might have met the second step proposed by Bahr.[7] However, due to the reduced sample size, we think more studies that re-evaluate the predictive performance of the model using data from new players are necessary.

5.5.3. Limitations

Although the model presented in this study shows moderate predictive scores, it should be acknowledged that more sophisticated algorithms (e.g. neural networks, genetic algorithms) might have developed models showing slightly better results than those found in the current study. However, the use of more complex algorithms would require that sport medicine practitioners must carry out complex math functions and operations, which might impact on the practical application of the model built dramatically. Thus, and in order to allow sport medicine practitioners to implement the model in their screening programmes, we decided to use decision trees algorithms as base classifiers because: a) they provide an easy to understand explanation for the classification result and can be used directly for decision making; b) they have been widely used to deal with imbalance data sets; and c) some of them are considered within the top-ten data mining algorithms.

The model developed in the present study was built with the goal of allowing sport medicine practitioners to accurately identify professional soccer and handball players at high risk of MUS_{INJ} during preseason screenings. To address this issue, we used several predictors (risk factors) as well as external (oversampling) and internal (ensembles) methods and a decision tree (ADTree) as base classifier in order to build a model with very good predictive accuracy. This set up allowed us to build a robust model (AUC score = 0.747; TPrate = 65.9; TNrate = 79.1) but also very complex in nature (black box approach). Therefore, although the model fulfils the goal for which it was built (make predictions), its complexity (10 different classifiers and 52 predictors) does not afford the opportunity to answer the question concerning why MUS_{INJ} happen.

Another potential limitation of the current study is the population used. The sport background of participants was professional soccer and handball players and the generalizability to other sport modalities and level of play cannot be ascertained. Therefore, future studies should be carried out in order to build predictive models to identify cohorts of players different than soccer and handball players at high risk of MUS_{INJ}. Finally, it should also be noted that the model is dependent of the predictors used in the training process and hence, practitioners must follow the same assessment methodologies used in the current study in order to replicate the current results and gain the applicability in their populations.

5.6. Conclusions

The current study is the first (to the best of our knowledge) that has used a model to identify professional soccer and handball players at high risk of MUS_{INJ} by applying a novel multifactorial approach and whose predictive ability has been determined through the exigent resampling technique called cross-validation. In this study the MUS_{INJ} risk model is comprised of 10 classifiers with a tree-shape structure and was developed thanks to the application of learning algorithms (on the training subsets) widely used in the data mining setting. Thus, the model reports an AUC score of 0.747 with true positive and negative rates of 65.9% and 79.1% respectively. We believe that the approach used here could replace the conventional statistical methods and can be used for coaches, physical trainers and medical practitioners to gain valuable information in the decision-making process aimed at reducing the number and severity of MUS_{INJ} in professional soccer and handball players.

5.7. Appendixes

Appendix 5.1. Description of the personal injury risk factors recorded.

Name	Labels
Sport	Soccer or handball
Player position	Goalkeeper, defender, midfielder or striker
Current level of play	1st division, 2nd B division, or 3rd division
Dominant leg	Right, left or two-footed
Age	Sub21, sub23, senior [23-30 y] or veteran [> 30y]
Body mass (kg)	< 71.65, 71.65-76.55, > 76.55-82.8 or > 82.8
Stature (cm)	< 1.76, 1.76-1.81, > 1.81-1.84 or > 1.84
BMI (kg/m^2)	< 22.75, 22.75-23.55, > 23.55-24.75 or > 24.75
History of MUS$_{INJ}$ last season	Yes or no

BMI: body mass index; MUS$_{INJ}$: Lower extremity muscle injury.

Appendix 5.2. Description of the psychological risk factors recorded.

Name	**Labels**
Sleep quality	< 3.5, 3.5-4.0 or > 4.0
Athlete Burnout Questionnaire	
a) Physical/emotional exhaustion	< 2.5 or ≥ 2.5
b) Reduced sense of accomplishment	≤ 2.5 or > 2.5
c) Sport devaluation	< 1.1, 1.1-1.49, > 1.49-1.9 or > 1.9

Appendix 5.3. Description of the measures obtained from the lower extremity range of motion assessment tests.

Name	Labels
PHF_{KF}	≤ 150 or > 150 (1)
PHF_{KE}	< 80, 80-100 or > 100 (2)
PHE	< 5, 5.0-15 or > 15 (5)
PHA	< 50, 50-70 or > 70 (3)
PHIR	< 45, 45-60 or > 60 (1)
PHER	< 40, 40-55 or > 55 (1)
PKF	< 110, 110-130 or > 130
ADF_{KE}	< 30, 30-40 or > 40 (5)
ADF_{KF}	< 30, 30-40 or > 40 (4)

ROM: range of motion; PHF_{KF}: passive hip flexion with knee flexed ROM; PHF_{KE}: passive hip flexion with knee extended ROM; PHE: passive hip extension ROM; PHA: passive hip abduction ROM; PHIR: passive hip internal rotation ROM; PHER: passive hip external rotation ROM; PKF: passive knee flexion ROM; ADF_{KE}: ankle dorsiflexion with knee extended ROM; ADF_{KF}: ankle dorsiflexion with knee flexed ROM.

(1): American Academy of Orthopaedic Association, 1965[225]; (2): Palmer & Epler, 2002[272]; (3): Gerhardt, 2002[237]; (4) Pope, Herbert & Kirwan (1998)[234]; (5) Cejudo, 2016.[224]

Appendix 5.4. Description of the measures obtained from the isokinetic knee flexion and extension strength assessment.

Measure	Labels	
	Dominant Leg	**Non-dominant Leg**
	Concentric Muscle Actions	
PT-KE$_{60}$	< 163.1, 163.1-184.605, > 184.605-211.05 or > 211.05	< 158.3, 158.3-179.14, > 179.14-197.3 or > 197.3
PT-KF$_{60}$	< 74.6, 74.6-87.505, > 87.505-104.65 or > 104.65	< 68.7, 68.7-84.9, > 84.9-98.2 or > 98.2
PT-KE$_{180}$	< 112.05, 112.05-129.3, > 129.3-146.3 or > 146.3	< 113.6, 113.6-128.495, > 128.495-146.55 or > 146.55
PT-KF$_{180}$	< 59.55, 59.55-70.4, > 70.4-81.4 or > 81.4-	< 60.1, 60.1-68.35, > 68.35-79.75 or > 79.75
PT-KE$_{240}$	< 98.05, 98.05-114.55, > 114.55-129.3 or > 129.3	< 95.45, 95.45-113.9, > 113.9-130.65 or > 130.65
PT-KF$_{240}$	< 57.8, 57.8-65.86, > 65.86-78.75 or > 78.75	< 55.7, 55.7-64.095, > 64.095-75.75 or > 75.75
PT-KE$_{300}$	< 90.75, 90.75-104.15, > 104.15-117.45 or > 117.45	< 85.45, 85.45-103.45, > 103.45-115.2 or > 115.2
PT-KF$_{300}$	< 54.55, 54.55-61.9, > 61.9-74.3 or > 74.3	< 48.2, 48.2-58.55, > 58.55-69.1 or > 69.1
APT-KE	< 45, 45-60 or > 60	
APT-KF	< 25, 25-35 or > 35	
	Eccentric Muscle Actions	
PT-KE$_{30}$	< 72.75, 72.75-90.105, > 90.105-109.15 or > 109.15	< 70.65, 70.65-84.12, > 84.12-95.75 or > 95.75
PT-KF$_{30}$	< 169.2, 169.2-207.42, > 207.42-242.2 or > 242.2	< 158.3, 158.3-198.1, > 198.1-236.9 or > 236.9
PT-KE$_{60}$	< 74.4, 74.4-91.14, > 91.14-109 or > 109	< 68.85, 68.85-86.3, 86.3-101.65 or > 101.65
PT-KF$_{60}$	< 175.6, 175.6-211.28, > 211.28-244.9 or > 244.9	< 156.3, 156.3-200.65, > 200.65-239.95 or > 239.95
PT-KE$_{180}$	< 73.6, 73.6-89.95, > 89.95-106 or > 106	< 68.5, 68.5-85.475, > 85.475-96.45 or > 96.45
PT-KF$_{180}$	< 155.35, 155.35-192.65, > 192.65-221.3 or > 221.3	< 157.2, 157.2-187.99, > 187.99-216.05 or > 216.05
APT-KE	< 25, 25-35 or > 35	
APT-KF	< 50, 50-65 or > 65	

Unilateral Conventional Ratios	
(1) KF/KE_{CONV60}	<0.47, 0.47-0.60 or >0.60
(2) $KF/KE_{CONV180}$	$\leq$ 0.60 or > 0.60
(3) $KF/KE_{CONV240}$	$\leq$ 0.60 or > 0.60
$KF/KE_{CONV300}$	< 0.6 0.6-0.8 or > 0.8
Unilateral Functional Ratios	
(4) KF/KE_{FUNC60}	< 0.6, 0.6-0.7 or > 0.7
$KF/KE_{FUNC180}$	$\leq$ 0.80 or > 0.80
(5) KF_{30}/KE_{240}	< 0.8, 0.8-1.0 or > 1.0
Bilateral Ratios	
KF/KF_{CON60}	No Asymmetry or Asymmetry
KF/KF_{CON180}	No Asymmetry or Asymmetry
KF/KF_{CON240}	No Asymmetry or Asymmetry
KE/KE_{CON60}	No Asymmetry or Asymmetry
KE/KE_{CON180}	No Asymmetry or Asymmetry
KE/KE_{CON240}	No Asymmetry or Asymmetry
KF/KF_{ECC60}	No Asymmetry or Asymmetry
KF/KF_{ECC180}	No Asymmetry or Asymmetry
KF/KF_{ECC240}	No Asymmetry or Asymmetry
KE/KE_{ECC60}	No Asymmetry or Asymmetry

APT: angle of peak torque; CON: concentric; CONV: conventional; FUNC: functional; ECC: eccentric; KE: knee extension; KF: knee flexion; PT: peak torque.
(1) Croisier et al. (2003)[273]; (2): Yeung et al. (2009)[274]; (3): Devan et al. (2004)[275]; (4): Dauty et al. (2003)[276]; (5) Croisier et al. (2002).[119]

Appendix 5.5. Description of the measures obtained from the dynamic postural control test.

Name	Labels	
	Dominant Leg	**Non-dominant Leg**
YBalance-Anterior	< 56.48, 56.48-60.055, > 60.055-63.86 or > 63.86	< 57.3, 57.3-60.895, > 60.895-65.27 or > 65.27
YBalance-Posteromedial	< 97.535, 97.535-104.055, > 104.055-108.885 or > 108.885	< 100.42, 100.42-104.905, > 104.905-108.8 or > 108.8
YBalance-Posterolateral	< 94.35, 94.35-99.485, > 99.485-106.79 or > 106.79	< 93.625, 93.625-99.175, > 99.175-104.48 or > 104.48
BilaRatio-YBalance-Anterior	No Asymmetry or Asymmetry	
BilaRatio-YBalance-Posteromedial	No Asymmetry or Asymmetry	
BilaRatio-YBalance-Posterolateral	No Asymmetry or Asymmetry	
YBalance-Composite	< 83.245, 83.245-87.86, > 87.86-92.035 or > 92.035	< 84.185, 84.185-87.985, > 87.985-91.84 or > 91.84

Bila: bilateral.

Appendix 5.6. Description of the measures obtained from the isometric hip abduction and adduction strength test.

Name	**Labels**	
	Dominant Leg	**Non-dominant Leg**
PT$_{ISOM}$-HipAbd	< 182.225, 182.225-204.09, > 204.09-221.17 or > 221.17	< 188.575, 188.575-208.9, > 208.9-227 or > 227
PT$_{ISOM}$-HipAbd-Normalised	< 2.39, 2.39-2.65, > 2.65-2.945 or > 2.945	< 2.485, 2.485-2.705, > 2.705-2.935 or > 2.935
PT$_{ISOM}$-HipAdd	< 187.75, 187.75-205.335, > 205.335-224.54 or > 224.54	< 181.975, 181.975-199.9, > 199.9-224.2 or > 224.2
PT$_{ISOM}$-HipAdd-Normalised	< 2.385, 2.385-2.735, > 2.735-2.99 or > 2.99	< 2.355, 2.355-2.655, > 2.655-2.945 or > 2.945
UnRatio-ISOM-HipAbd/HipAdd	<0.936, 0.936-1.045, > 1.045-1.17 or > 1.17	<0.905, 0.905-0.973, >0.973.065 or > 1.065
BilaRatio-PT$_{ISOM}$-HipAbd/HipAdd	No Asymmetry or Asymmetry	

Abd: abduction; Add: adduction; Bila: bilateral; ISOM: isometric; PT: peak torque; Uni: unilateral.

Appendix 5.7. Description of the measures obtained from the core stability test.

Name	Labels
CS_{NF}	< 4.895, 4.895-6.14, > 6.14-7.83 or > 7.83
CS_{WF}	< 4.335, 4.335-5.475, > 5.475-6.84 or > 6.84
CS_{ML}	< 6.915, 6.915-8.47, > 8.47-9.62 or > 9.62
CS_{AP}	< 7.19, 7.19-8.33, > 8.33-9.865 or > 9.865
CS_{CD}	< 9.01, 9.01-10.555, > 10.555-12.375 or > 12.375

CS_{NF}: unstable sitting without feedback; CS_{WF}: unstable sitting with feedback; CS_{ML}: unstable sitting while performing medial-lateral displacements with feedback; CS_{AP}: unstable sitting while performing anterior-posterior displacements with feedback; CS_{CD}: unstable sitting while performing circular displacements with feedback.

Appendix 5.8. Algorithms used in the data processing phase.

Base classifiers

Abbr.	*Method*	*Short Description*
J48	J48	Algorithm for generating a pruned or unpruned C4.5 decision tree
SCart	SimpleCart	Algorithm for implementing minimal cost-complexity pruning
ADTree	ADTree	Alternating decision tree
RTree	RandomTree	Algorithm that considers K randomly chosen attributes at each node of the tree

Resampling techniques

Abbr.	*Method*	*Short Description*
SMT	SMOTE	Each decision tree applied on data set previously pre-processed with Smote
ROS	Random over sampling	Each decision tree applied on data set previously pre-processed with random over sampling
RUS	Random under sampling	Each decision tree applied on data set previously pre-processed with random under sampling

Classis Ensembles

Abbr.	*Method*	*Short Description*
ADAB	AdaBoost	Classic AdaBoost, without using confidences
M1	AdaBoost.M1	Multi-class AdaBoost, slightly different weight update
BAG	Bagging	Classic Bagging, resampling with replacement, bag size equal to original data set size.

Boosting-based Ensembles

Abbr.	*Method*	*Short Description*
SBO	SmoteBoost	AdaBoost.M2 with Smote in each iteration
RUS	RusBoost	AdaBoost.M2 with random undersampling in each iteration

Cost-sensitive learning

Abbr.	*Method*	*Short Description*
MetaCost	MetaCost	Makes base classifier cost-sensitive by passing it to Bagging
CS-Classifier	Cost Sensitive Classifier	Makes base classifier cost-sensitive.

Bagging-based Ensembles

Abbr.	*Method*	*Short Description*
OBAG	OverBagging	Bagging with oversampling of the minority class.
UBAG	Underbagging	Bagging with undersampling of the majority class.
SBAG	SmoteBagging	Bagging where each bag´s Smote quantity varies

Ensembles with a cost-sensitive based classifier

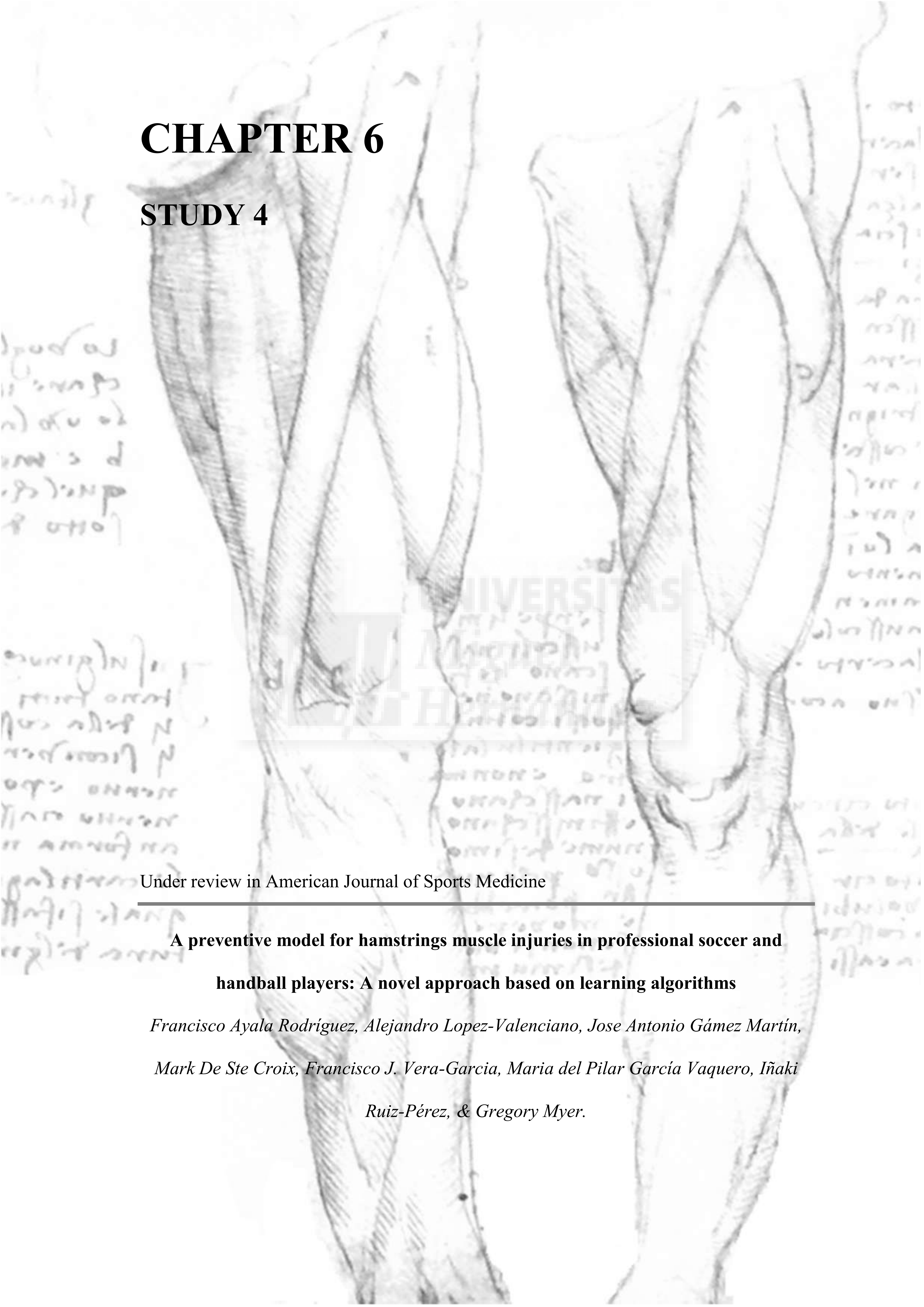
Abbr.	*Method*	*Short Description*
CS-SBAG	Cost sensitive SmoteBagging	SmoteBagging with an asymmetric classification cost matrix in the base classifier
CS.OBAG	Cost sensitive OverBagging	OverBagging with an asymmetric classification cost matrix in the base classifier
CS- UBAG	Cost sensitive UnderBagging	UnderBagging with an asymmetric classification cost matrix in the base classifier

Appendix 5.9. Risk factor measures included in the model for predicting muscle injuries and the number of times that they appear in the classifiers.

Risk Factor	Nº of Classifiers
Personal measures	
Age group	1
History of MUS$_{INJ}$ last season	4
Maximal level of play achieved	2
BMI	1
Psychological measures	
Sleep Quality	1
Sport Devaluation	5
Dynamic postural control measures	
YBalance-Anterior- Dominant Leg	1
YBalance-Anterior-Non-dominant Leg	2
YBalance-Composite-Dominant Leg	1
YBalance-Posterolateral-Non-dominant Leg	1
YBalance-Posteromedial-Non-dominant Leg	1
BilaRatio-YBalance-Posterolateral	1
Isometric hip abduction and adduction strength measures	
BilaRatio-PT$_{ISOM}$-HipAdd	1
PT$_{ISOM}$-HipAdd-Dominant Leg	2
PT$_{ISOM}$-HipAdd-Non-dominant	1
UniRatio-PT$_{ISOM}$-HipAbd/HipAdd	1
Lower extremity joint range of motion measures	
ROM-ADF$_{KF}$-Non-dominant Leg	1
ROM-PHF$_{KE}$-Dominant Leg	1
ROM-PKF-Dominant Leg	1
ROM-PKF-Non-dominant Leg	3
Core stability measures	
Core- CS$_{NF}$	1
Core- CS$_{WF}$	1
Core- CS$_{CD}$	1
Isokinetic knee flexion and extension strength measures	
APT-KE$_{CON240º/s}$-Dominant leg	2
APT-KE$_{CON240º/s}$-Non-dominant Leg	1
APT-KE$_{CON60º/s}$-Dominant leg	2
APT-KE$_{CON60º/s}$-Non-dominant leg	1
APT-KE$_{ECC180º/s}$-Dominant Leg	3
APT-KE$_{ECC60º/s}$-Dominant leg	1
APT-KF$_{CON180º/s}$-Dominant Leg	2
APT-KF$_{CON60º/s}$-Dominant Leg	3
APT-KF$_{CON60º/s}$-Non-dominant Leg	1
APT-KF$_{ECC30º/s}$-Dominant Leg	2
APT-KF$_{ECC60º/s}$-Non-dominant Leg	2
BilaRatio-KF$_{CON180º/s}$	1
BilaRatio-KF$_{CON240º/s}$	1
BilaRatio-KF$_{ECC240º/s}$	2
PT-KE$_{CON180º/s}$-Non-dominant Leg	1
PT-KE$_{CON240º/s}$-Non-dominant Leg	3
PT-KE$_{CON300º/s}$-Dominant Leg	2
PT-KE$_{CON300º/s}$-Non-dominant Leg	1
PT-KE$_{CON60º/s}$-Non-dominant Leg	1
PT-KE$_{ECC180º/s}$-Non-dominant Leg	1

PT-KF$_{CON180º/s}$-Dominant Leg	1
PT-KF$_{CON240º/s}$- Dominant	1
PT-KF$_{CON240º/s}$-Non-dominant Leg	1
PT-KF$_{CON300º/s}$-Dominant Leg	4
PT-KF$_{CON60º/s}$-Non-dominant Leg	2
PT-KF$_{ECC180º/s}$-Non-dominant Leg	1
PT-KF$_{ECC30º/s}$-Non-dominant Leg	3
PT-KF$_{ECC60º/s}$-Non-dominant Leg	3
UnilRatio KF/KE$_{CON60º/s}$-Dominant Leg	3
UniRatio-KF/KE$_{CON240}$-Dominant Leg	1

Abd: abduction; Add: adduction; ADF$_{KF}$: ankle dorsiflexion with knee flexed ROM; APT: angle of peak torque; Bila: bilateral; BMI: body mass index; CON: concentric; CS$_{CD}$: unstable sitting while performing circular displacements with feedback; CS$_{NF}$: unstable sitting without feedback; CS$_{WF}$: unstable sitting with feedback; ISOM: Isometric; ECC: eccentric; KE: knee extension; KF: knee flexion; MUS$_{IN}$: Muscle injury; PHF$_{KE}$: passive hip flexion with knee extended ROM; PKF: passive knee flexion ROM; PT: peak torque; ROM: range of motion; s: seconds; Uni: unilateral; º: degree.



CHAPTER 6

STUDY 4

Under review in American Journal of Sports Medicine

A preventive model for hamstrings muscle injuries in professional soccer and handball players: A novel approach based on learning algorithms

Francisco Ayala Rodríguez, Alejandro Lopez-Valenciano, Jose Antonio Gámez Martín, Mark De Ste Croix, Francisco J. Vera-Garcia, Maria del Pilar García Vaquero, Iñaki Ruiz-Pérez, & Gregory Myer.

CHAPTER 6

STUDY 4

A preventive model for hamstrings muscle injuries in professional soccer and handball players: A novel approach based on learning algorithms

by

Ayala F, Lopez-Valenciano A, Gámez JA, De Ste Croix M, Vera-Garcia FJ, García-Vaquero MP, Ruiz-Pérez I, & Myer G.

6.1. Abstract

Background: Hamstring strain injury (HSI) is one of the most prevalent injuries reported in professional sport despite the substantive effort made by the scientific community and medical practitioners. The application of contemporary statistical approaches coming from machine learning and data mining environments, which develop more robust predictive models to identify players at high risk of HSI, might support injury prevention strategies of the future.

Purpose: To analyse and compare the predictive ability of a range of decision tree algorithms in order to select the best performing injury risk factor model to identify professional players at high risk of HSIs.

Study Design: Prospective cohort study (level of evidence 2).

Methods: A total of 132 male professional soccer (n = 98) and handball (n = 34) players underwent a preseason screening evaluation of a number of individual, psychological and neuromuscular measures. Furthermore, injury surveillance was employed to capture all the HSI occurring in the 2013/2014 seasons. The predictive ability (determined through the area under the receiver operating characteristic curve [AUC]) of several models built by applying a range of learning techniques were analysed and compared in order to find the best model for identifying risk of HSI.

Results: There were 21 HSIs over the follow up period. Injury distribution between the legs was 57.2% dominant leg and 42.8% non-dominant leg. The model generated by the SmooteBoostM1 technique with a cost-sensitive alternating decision tree as base classifier reported the best evaluation criteria (AUC score = 0.867) and hence was considered the best for predicting HSI.

Conclusions: The prediction model showed high accuracy for identifying professional soccer and handball players at risk of HSI during preseason screenings. Therefore, the model developed might help coaches, physical trainers and medical practitioners in the decision-making process for injury prevention.

Clinical Relevance: Clinicians and sport medical practitioners might use the model built to identify players at high risk of hamstring injuries during the preseason screenings and thus it would help in the decision-making process for injury prevention.

What is known about the subject: Despite some risk factors (e.g. strength imbalances, older age, poor flexibility, previous injury) have demonstrated a strong relationship with hamstrings injuries, the ability of the cut-off scores proposed to predict injuries is not acceptable for screening purposes.

What this study adds to existing knowledge: This study has developed a model to identify professional soccer and handball players at high risk of hamstring injuries by applying a novel contemporary statistical approach coming from machine learning environments and whose predictive ability has been determined through the exigent resampling technique called cross-validation

Keywords: *injury prevention, learning algorithms, modelling, screening, decision-making.*

6.2. Introduction

The Hamstring strain injury (HSI) is the most prevalent injury reported in professional sport, with HSIs alone accounting for between 6% and 29% of all injuries in Australian Rules soccer,[277] rugby union,[278,279] soccer,[41] basketball[280] and track sprinters.[281] The frustrating issue of HSIs for the player is not only explained by the high prevalence of these injuries,[48] but also by; a) the prolonged duration of symptoms (resulting in a mean of 14-22 training and match days lost per injury)[41]; b) poor healing responses; c) a high risk of re-injury rate of 12-31%[121,278,282-284]; and c) their overall burden could be extremely significant for clubs.[159]

Prior to establishing injury prevention programmes, it is essential to identify players at high risk of HSI. Several prospective studies have identified a number of modifiable (e.g. strength, joint ranges of motion, core stability, etc.) and non-modifiable (e.g. age, sex, history of HSI, etc.) risk factors that have demonstrated a statistically significant relationship with HSIs.[20,23,93,116-125,207,285] Although the presence of a statistically significant association indicates that there may be a causal relationship between the factor and injury incidence, this knowledge is likely insufficient to identify players at high risk of HSIs.[7] Accordingly, some studies have defined markers or cut-off scores for specific risk factors in an attempt to identify players at high and low risk of HSI.[93,118,119,121,122,207]

However, despite a substantive effort made in recent years by the scientific community and medical practitioners to firstly identify players at high risk of HSIs and then apply tailored injury prevention programmes, recent evidence has demonstrated that HSI incidence has not decreased, but has increased slightly over recent years.[48,286]

Two different arguments appear to be behind the lack of generality of the proposed cut-off scores and that could explain why they do not permit the identification of players at high risk. Firstly, the generality (external validity) of the cut-off scores proposed for certain injury risk factors (e.g. strength imbalance, joint range of motion [ROM]) might be limited since their predictive abilities to identify new athletes at high risk of HSI have not been verified in a new population of athletes, different that the one used for defining them (e.g. cross validation).[7,115] This suggests that cut-off scores might be overfitted (i.e. their predictive ability is adjusted to the data set used in their learning process), which will give overly optimistic results and hence, they may not be acceptable for screening purposes. This appears to be supported by the fact that the cut-off scores defined by some prospective studies (mainly those related to strength measures) have not been later ratified by others using similar designs and assessment methodologies but with different samples of athletes.[20,23,93,116-125] For example, while Croisier et al.,[118] and Dauty et al.,[120] found that professional soccer players with reciprocal (functional) hamstring-to-quadriceps ratios (H/Q) lower than 0.8 were at higher risk of sustaining a HSI, van Dyk et al.,[125] did not identify this strength ratio measure as a risk of HSI. The second issue with the current evidence base is that the above-cited prospective studies have identified potential risk factors for HSI according to the presence of statistically significant relationships (based on odds ratios, certain values of p statistic [mainly $p < 0.05$]) with HSI. However, based on the general agreement that the aetiology of HSIs is multifactorial and that some relationships of conditional dependence might exist among factors,[126] it is possible that the influence of an specific factor on the likelihood of suffering a HSI might not be statistically significant ($p < 0.05$) in itself, but relevant when it is used in conjunction with several other factors to develop a more robust predictive model. In other words, combining information from several modifiable and non-modifiable risk factors might lead to the development of a more robust model with an improved predictive ability.

The application of contemporary statistical approaches (e.g. supervised learning algorithms) derived from machine learning and data mining environments, which have been specifically designed to deal with problems where a large number of factors are involved and the use of resampling techniques (e.g. cross-validation, bootstrap and leave-one-out) may overcome the limitations inherent to the current body of knowledge and give light to better identify players at high risk of HIS.[127] In particular, decision tree algorithms are powerful statistical tools for prediction that have been used in several medical diagnosis studies reporting excellent results.[135-137] These learning algorithms have several advantages compared to traditional approaches (e.g. logistic regression)[138]: a) they simplify complex relationships between input variables and target variables by dividing the original input variables into significant subgroup; b) they are easy to understand and interpret; c) they use non-parametric approach without distributional assumptions; d) it is easy to handle missing values without needing to resort to imputation; e) it is easy to handle heavy skewed data without needing to resort to data transformation; and f) they are robust to outliers. In addition, decision tree algorithms are also able to select only those variables that are considered necessary to develop robust predictive models, reducing the number of variables necessary for a decision-making process. They can also be used as based classifiers in ensemble methods (also known as multiple classifier systems)[139] to obtain better predictive performance scores. Finally, the implementation of resampling techniques (e.g. cross validation, which involves [repeated multiple times] that a subset of the sample [training subset] is used to fit a model and the remaining subset [test subset] is used to estimate the efficacy of the model) to validate the model might provide a more accurate estimation of the predictive performance and increase its generality.[115]

Therefore, the main purpose of this study was to analyse and compare the predictive ability of a range of decision tree algorithms in order to select the best performing injury risk factor model to identify professional players at high risk of HSIs.

6.3. Methods

6.3.1. Participants

A total of 132 male professional soccer (n = 98) and handball (n = 34) players took part in the current study. Soccer players were recruited from four different soccer teams that were engaged in the 1st (one team, n = 25) and 2nd B (three teams, n = 73) Spanish National Soccer League divisions. Handball players were recruited from three different handball teams that were engaged in the 1st (one team, n = 11) and 3rd (two teams, n = 23) National Handball League divisions.

The exclusion criteria were: a) presence of orthopaedic problems that prevented participants from executing properly of one or more of the neuromuscular tests selected for this study; and b) players who were transferred to other clubs and did not finish the 9-month follow up period. Only new injuries we used for any player sustaining multiple HSIs.

Prior to study participation, experimental procedures and potential risks were fully explained to the participants in verbal and written form, and written informed consent was obtained from them. An Institutional Research Ethics committee approved the study protocol prior to data collection, conforming to the recommendations of the Declaration of Helsinki.

6.3.2. Study design

A prospective cohort design was used to address the purposes of this study. In particular, all the HSI accounted for within the 9 months (2013/2014 season) following the initial testing session were prospectively collected for all players.

Players underwent a preseason evaluation of a number of personal, psychological and neuromuscular measures, most of them considered potential sport-related injury risk factors. In each soccer and handball team, the testing session was conducted at the middle-end of the preseason phase of the year.

6.3.3. Testing procedure

The testing session was divided into three different parts (Figure 6.1). The first part of the test session was used to obtain information related to the participants' personal or individual characteristics. The second part was designed to assess psychological measures related to sleep quality and athlete burnout. Finally, the third part of the session was used to assess a number of neuromuscular measures.

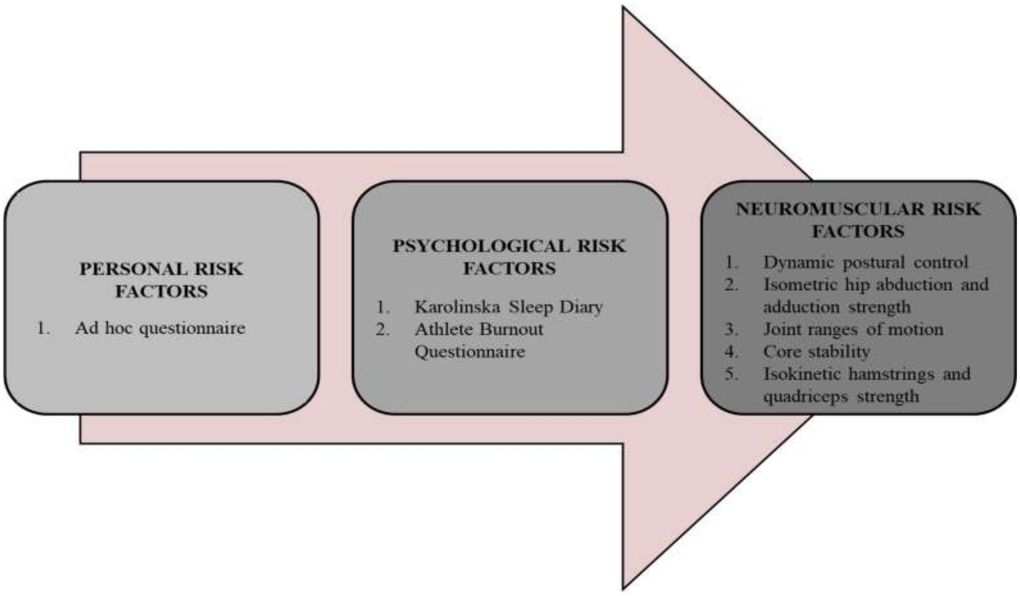


Figure 6.1. Testing procedure.

6.3.3.1. Personal or individual risk factors

The ad hoc questionnaire designed by Olmedilla et al.,[255] was used to record personal or individual features that have been defined as potential non-modifiable risk factors for sport injuries. Through this questionnaire sport-related background (sport, player position, current level of play, dominant leg [defined as the participant´s kicking leg]) and demographic (age, body mass and stature) features were recorded. In addition, the presence within the last season (yes or no) of HSIs with a total time taken to resume full training and match > 8 days was also recorded (self-reported). Table 6.1 displays a description of all the personal risk factors recorded.

Table 6.1. Description of the personal injury risk factors recorded.

Name	Labels
Sport	Soccer or handball
Player position	Goalkeeper, defender, midfielder or striker
Current level of play	1st division, 2nd B division, or 3rd division
Dominant leg	Right, left or two-footed
Age	Sub21, sub23, senior [23-30 y] or veteran [> 30y]
Body mass (kg)	< 73.65, 73.75-80.3 or > 80.3
Stature (cm)	< 1.785, 1.785-1.835 or > 1.835
History of HSI last season	Yes or no

cm: centimetre; HSI: hamstring strain injury; kg: kilogram.

6.3.3.2. Psychological risk factors

Sleep quality and athlete burnout variables were measured through two validated and worldwide used likert scales. The Spanish version of the Pittsburgh Sleep Diary[256] was used to measure the sleep quality of the soccer and handball players.

The Spanish version of the Athlete Burnout Questionnaire[257] was used to assess the three different dimensions that comprise athlete burnout: a) physical/emotional exhaustion; b) reduced sense of accomplishment; and c) sport devaluation. Appendix 6.1 displays a description of all the psychological risk factors recorded.

6.3.3.3. Neuromuscular risk factors

Prior to the neuromuscular risk factor assessment, all participants performed the dynamic warm-up designed by Taylor et al.[230] The overall duration of the entire warm-up was approximately 15-20 min. The assessment of the neuromuscular risk factors started 3-5 min after the dynamic warm-up.

In the experimental session, participants were assessed from a number of neuromuscular performance measures obtained from 5 different testing manoeuvres: 1) dynamic postural control[258]; 2) isometric hip abduction and adduction strength[259]; 3) lower extremity joint ranges of motion[287]; 4) core stability[260]; and 5) isokinetic hamstrings and quadriceps strength[262]. For a matter of space, the testing manoeuvres are not described below and the reader is to refer to their original sources. Furthermore, appendixes 6.2-6.6 display a description of all the neuromuscular risk factors recorded.

The order of the tests was consistent for all participants (Figure 6.1) and was established with the intention of minimizing any possible negative influence among variables. A 5-min rest interval was given between consecutive testing manoeuvres.

6.3.4. Injury Surveillance

Similar to previous studies[122,125] and following the recommendations made by the International Injury Consensus Group,[19] a HSI was defined as acute pain in the muscle location that occurred during training or match and resulted in the immediate termination of play and inability to participate in the next training session or match.

The club medical staff of each club recorded HSIs on an injury form that was sent to the study group each month. For all HSIs that satisfied the inclusion criteria, team medical staff provided the following details to investigators: leg injured (dominant/non-dominant), injury severity based on lay off time from soccer or handball (slight/minimal [0-3 days], mild [4-7 days], moderate [8-28 days], and severe [> 28 days]), date of injury, moment (training or match), whether it was a recurrence (defined as an HSIs that occurred in the same leg and during the same season as the initial injury), and total time taken to resume full training and match. At the conclusion of the 9 months follow up period, all data from the individual clubs were collated into a central database, and discrepancies were identified and followed up at the different clubs to be resolved. Some discrepancies among medical staff teams were found to diagnose minimal HSIs and to record their total time lost. To resolve these inconsistencies in the injury surveillance process (risk of misclassification of the players), only HSIs showing a time lost > 4 days (minor to severe) were selected for the subsequent statistical analysis.

6.3.5. Statistical analysis

The statistical analysis framework carried out in this study for analysing and comparing the behaviours of several machine learning techniques with the aim of finding the best model for predicting HSIs in professional soccer and handball players was based on a supervised learning perspective. From a statistical standpoint, the problem can be stated as follows: given a set of features F (in our case risk factors) and a target (discrete) variable (in our case HIS [yes or no]), named class, C, we want to estimate/learn a map function $M:F \rightarrow C$. Thus, the statistical analysis comprised two stages:

1. Data pre-processing. At this stage, the data set was prepared to apply the machine learning techniques. To optimise this aspect, pre-processing methods such as data cleaning and data discretization were applied.
2. Data processing. At this stage, the most powerful techniques reported by Galar et al., [113] and Elkarami et al.,[263] to address learning with imbalanced data sets were applied in order to build models for predicting HSIs. In particular, a study on the performance of some proposals for pre-processing, cost-sensitive learning and ensemble-based methods was carried out. Three classic decision tree algorithms were used as base classifiers in each method: J48,[265] ADTree[267] and SimpleCart.[266] The model with the best performance scores (determined through a fivefold stratified cross validation [SCV] technique) was finally selected to predict HSI in professional soccer and handball players.

A complete description of the statistical techniques carried out in both stages, data pre-processing and data processing, has been written in the Appendix 6.7.

6.4. Results

6.4.1. Hamstring strain injuries epidemiology

There were 21 HSI over the follow up period and all of them were used to train the models. Injury distribution between the legs was 57.2% dominant leg and 42.8% non-dominant leg. In term of severity, most of injures were categorized as moderate (n = 17) while only 4 cases were considered minor and no severe injuries were recorded.

6.4.2. Predictive model for lower extremity muscle injuries

Table 6.2 shows the average area under the receiver operating characteristic curve (AUC), true positive rate (TPrate) and true negative rate (TNrate) results for all oversampling and ensemble learning methods separately for each decision tree base classifier.

Table 6.2. Average area under the receiver operating characteristic curve, true positive rate and true negative rate results for all the decision tree methodologies in isolation and after having been applied in them the oversampling and ensemble techniques selected.

Technique	AUC	TPrate	TNrate
Cost-sensitive base classifiers			
J48	**0.787**	**76.2**	**79.8**
ADTree	0.789	57.1	0.768
Scart	0.675	66.7	67.7
Oversampling techniques			
SMT			
J48	0.694	66.7	73.7
ADTree	**0.739**	**47.6**	**84.6**
Scart	0.700	61.9	75.8
Boosting-based Ensembles			
SBOM1			
J48	0.671	42.9	92.9
ADTree	0.867	71.4	89.9
Scart	0.621	61.9	72.7
RUSB			
J48	0.783	42.9	87.9
ADTree	**0.859**	**66.7**	**89.9**
Scart	0.576	23.8	86.9
Bagging-based Ensembles			
OB			
J48	0.802	47.6	92.9
ADTree	**0.871**	**71.4**	**84.8**
Scart	0.767	57.1	80.8
SBAG			
J48	0.766	33.3	90.9
ADTree	**0.892**	**47.6**	**93.9**
Scart	0.778	61.9	76.8

Abbreviations can be found in appendix 5.8. Highlighted in bold is the method that obtained the best performing result within each method. Highlighted in grey is the model considered as the best for predicting hamstring strain injury.

The ADTree base classifier reported the best performance in most of the methods analysed. In fact, the final model was built using the SMOTEBoostM1 ensemble method with the ADTree as the base classifier using a reweighted training instance (cost-sensitive) approach.

Therefore, the final model selected to predict HSI in professional soccer and handball players was comprised by 10 different cost sensitive ADTress classifiers (Figures 6.2-6.11). The cost matrix for cost-sensitive classifier was set to $C\left\{\frac{0}{1}\middle|\frac{10}{0}\right\}$.

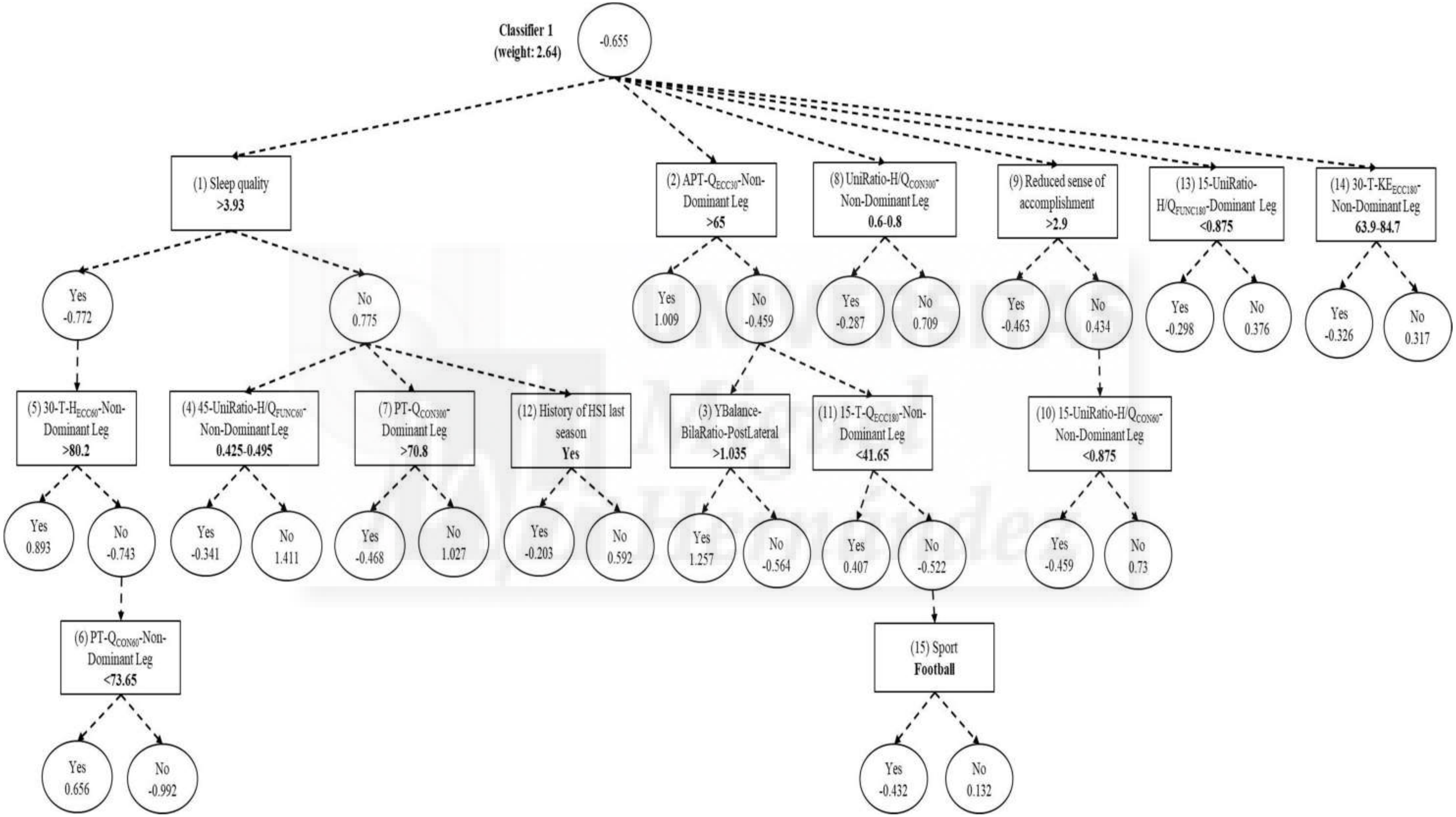


Figure 6.2. Hamstring strain injury predictive model, classifier 1.

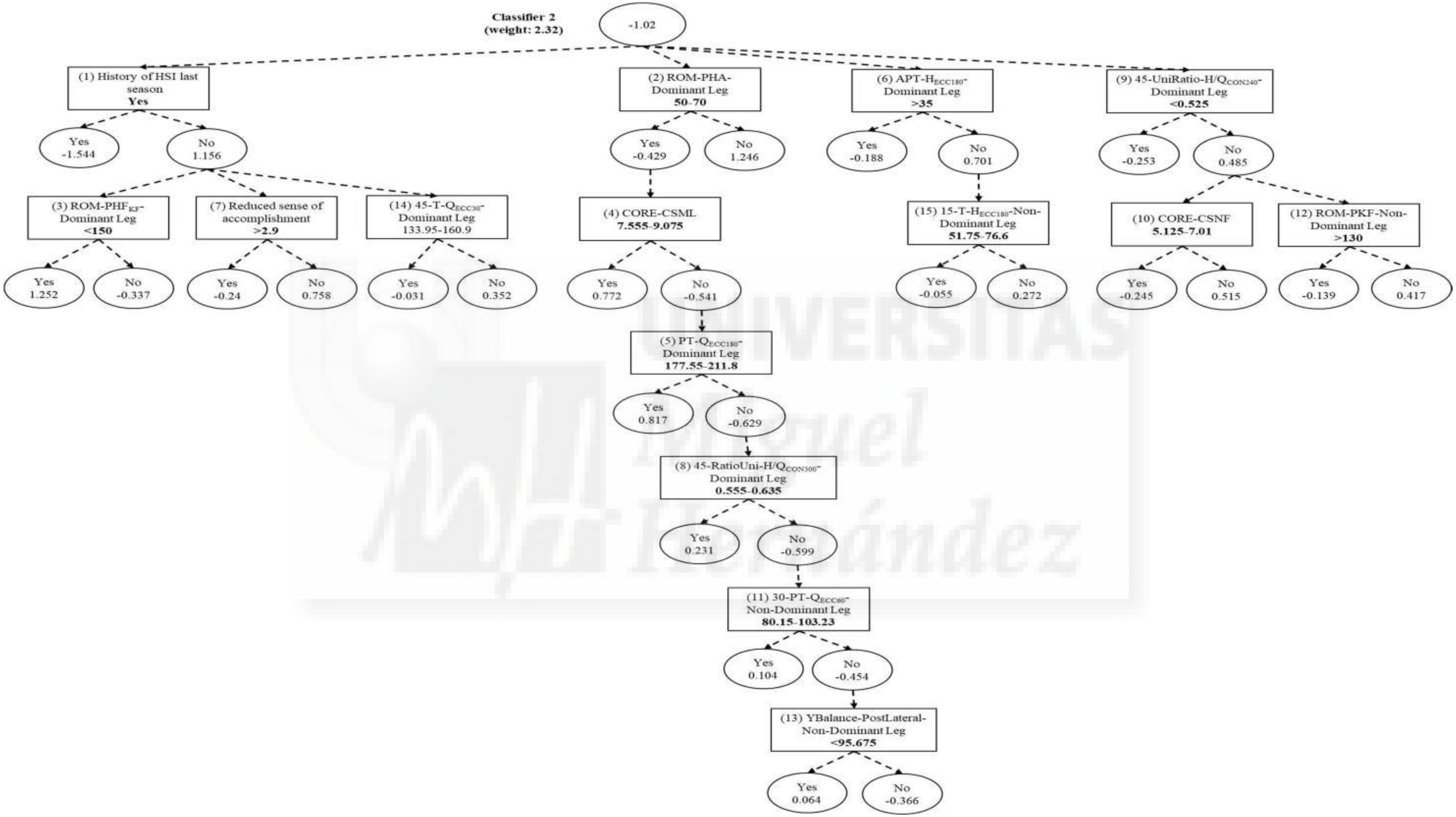


Figure 6.3. Hamstring strain injury predictive model, classifier 2.

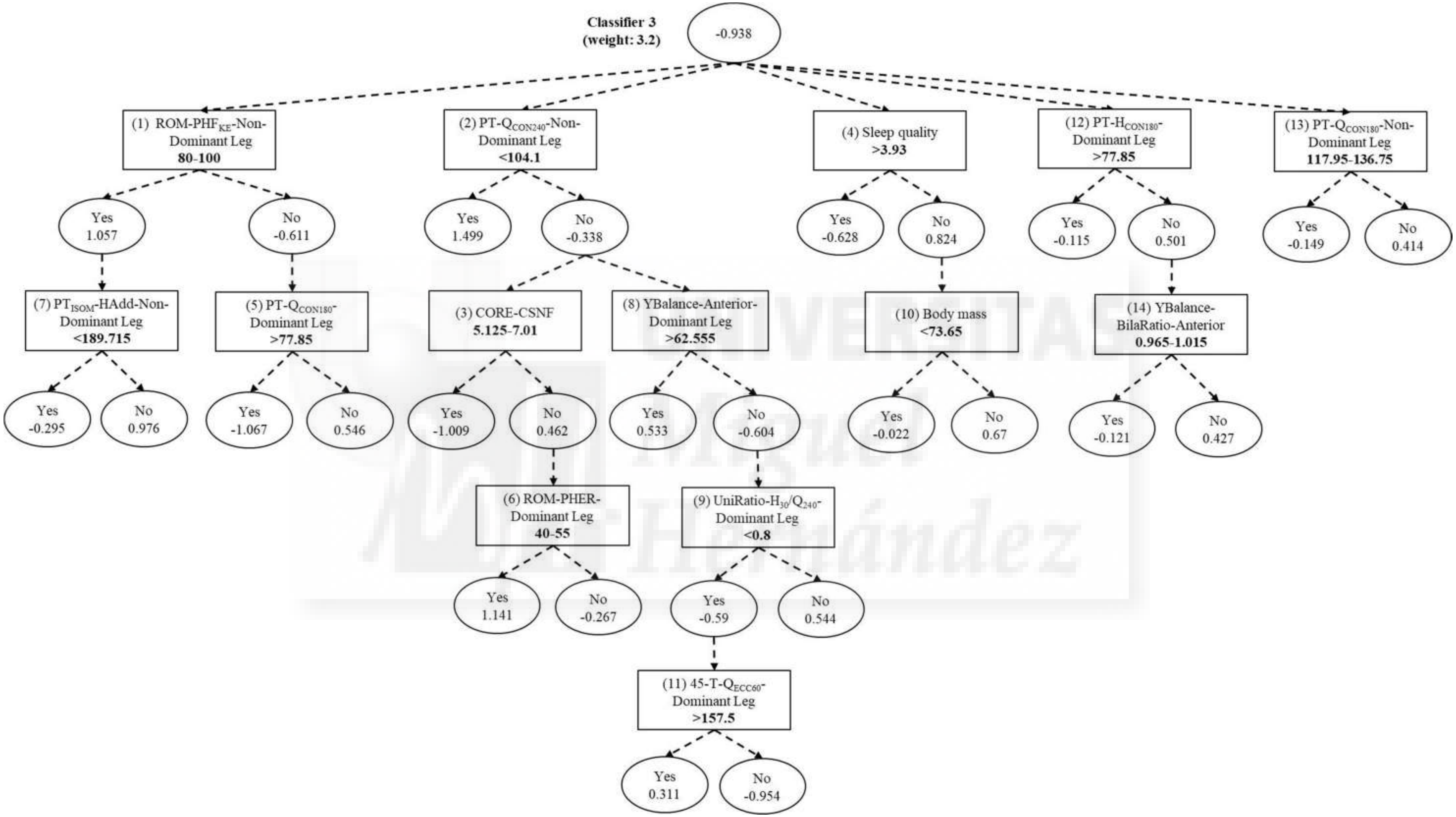


Figure 6.4. Hamstring strain injury predictive model, classifier 3.

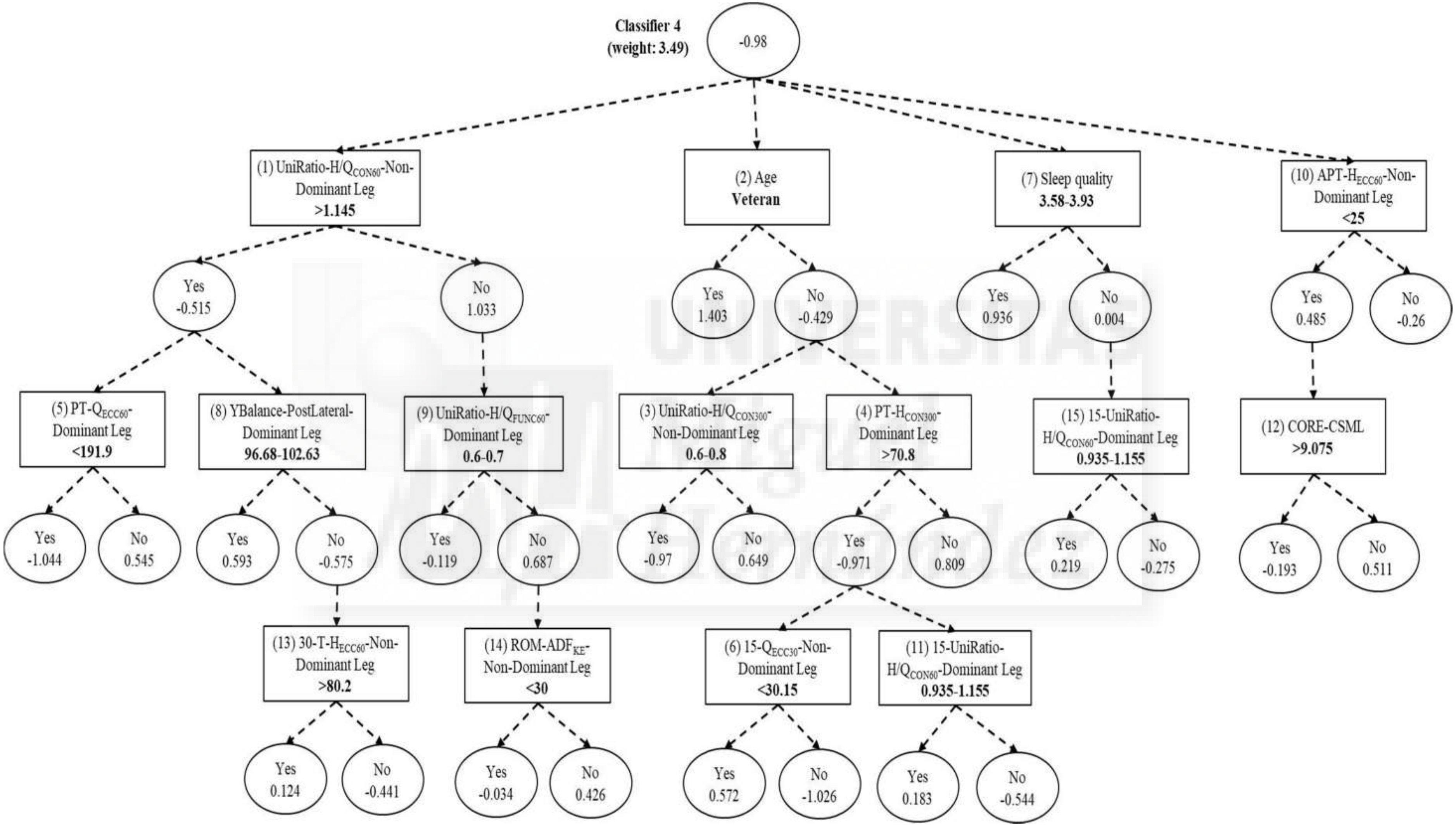


Figure 6.5. Hamstring strain injury predictive model, classifier 4.

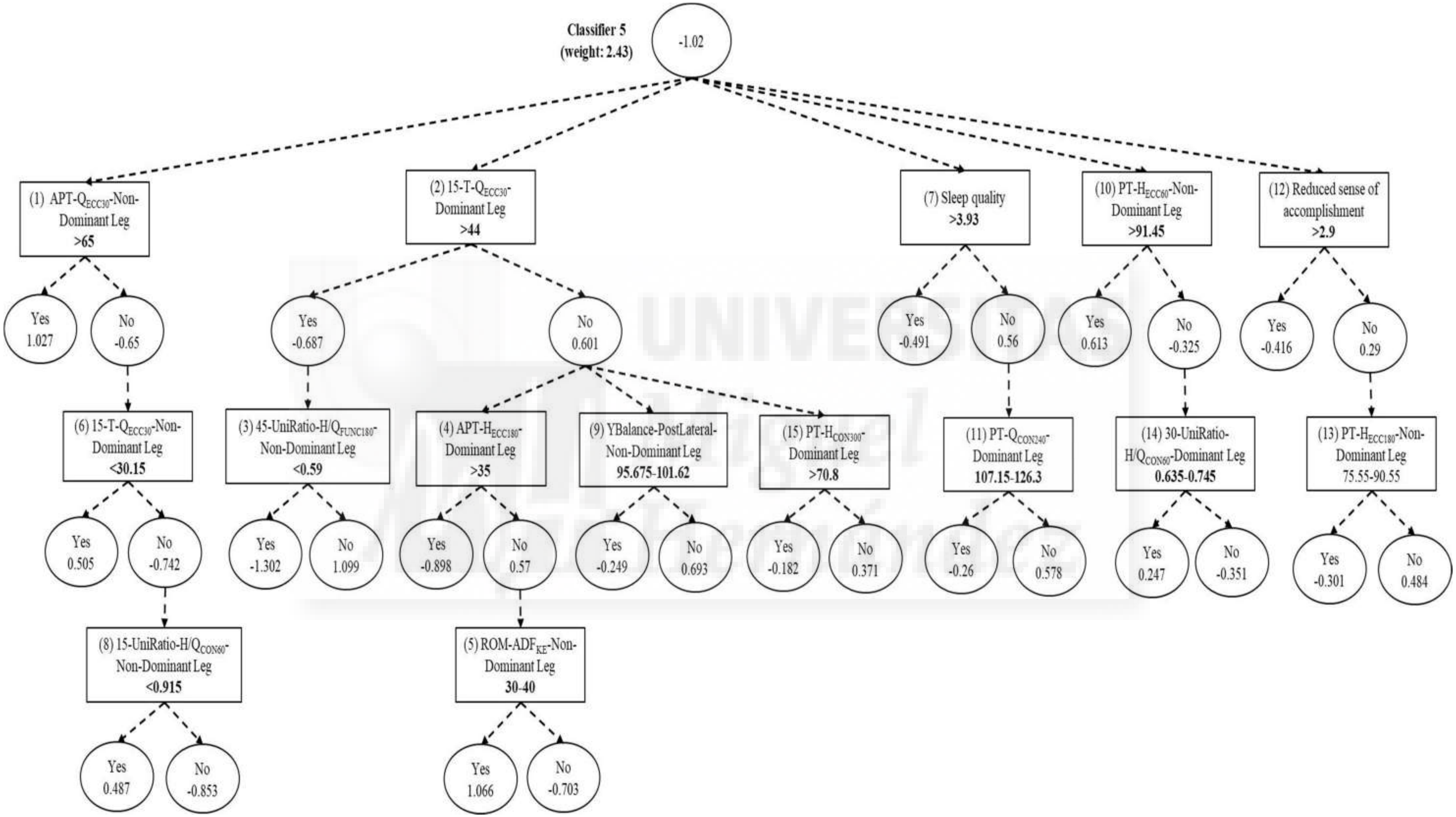


Figure 6.6. Hamstring strain injury predictive model, classifier 5.

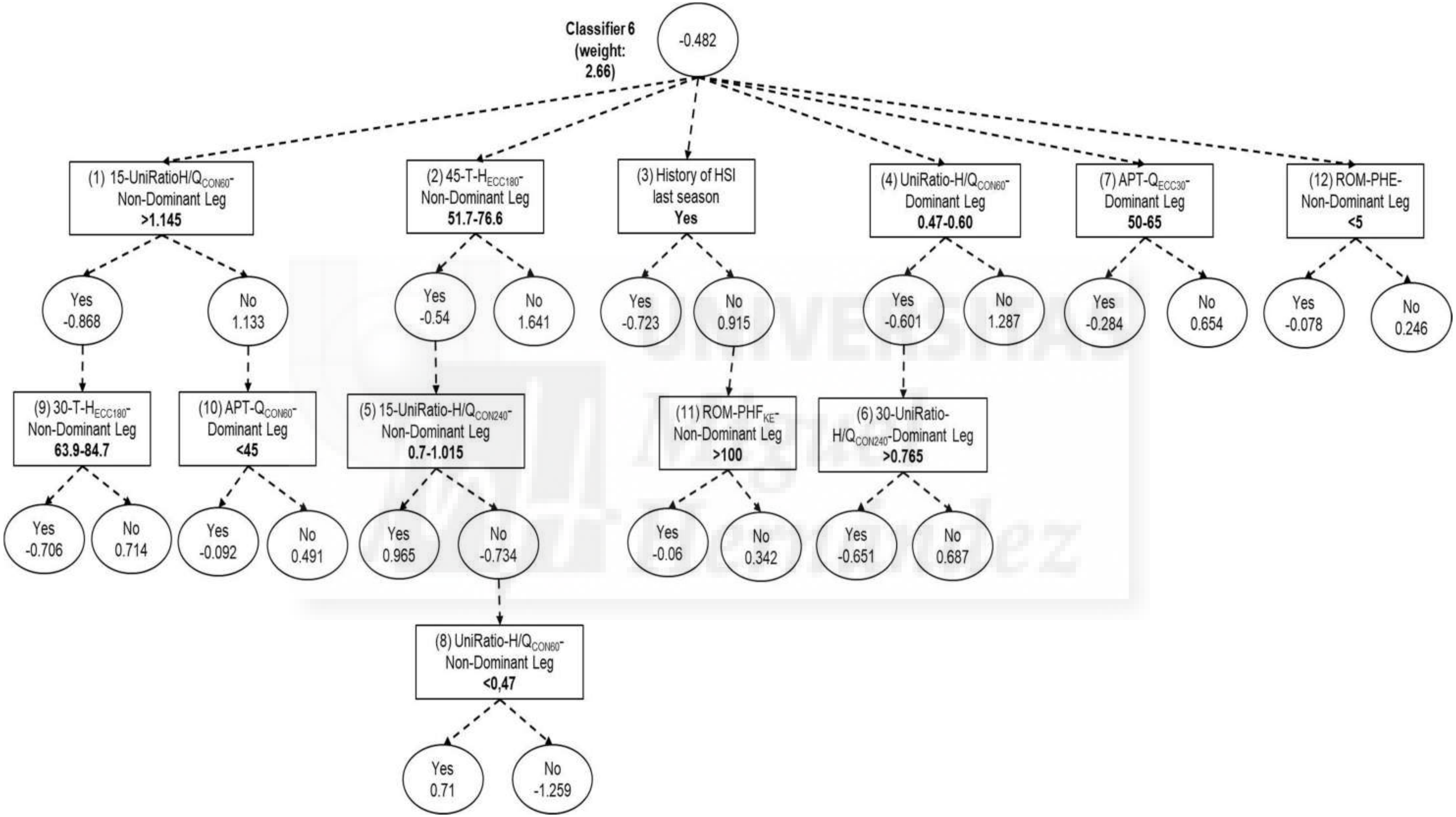


Figure 6.7. Hamstring strain injury predictive model, classifier 6.

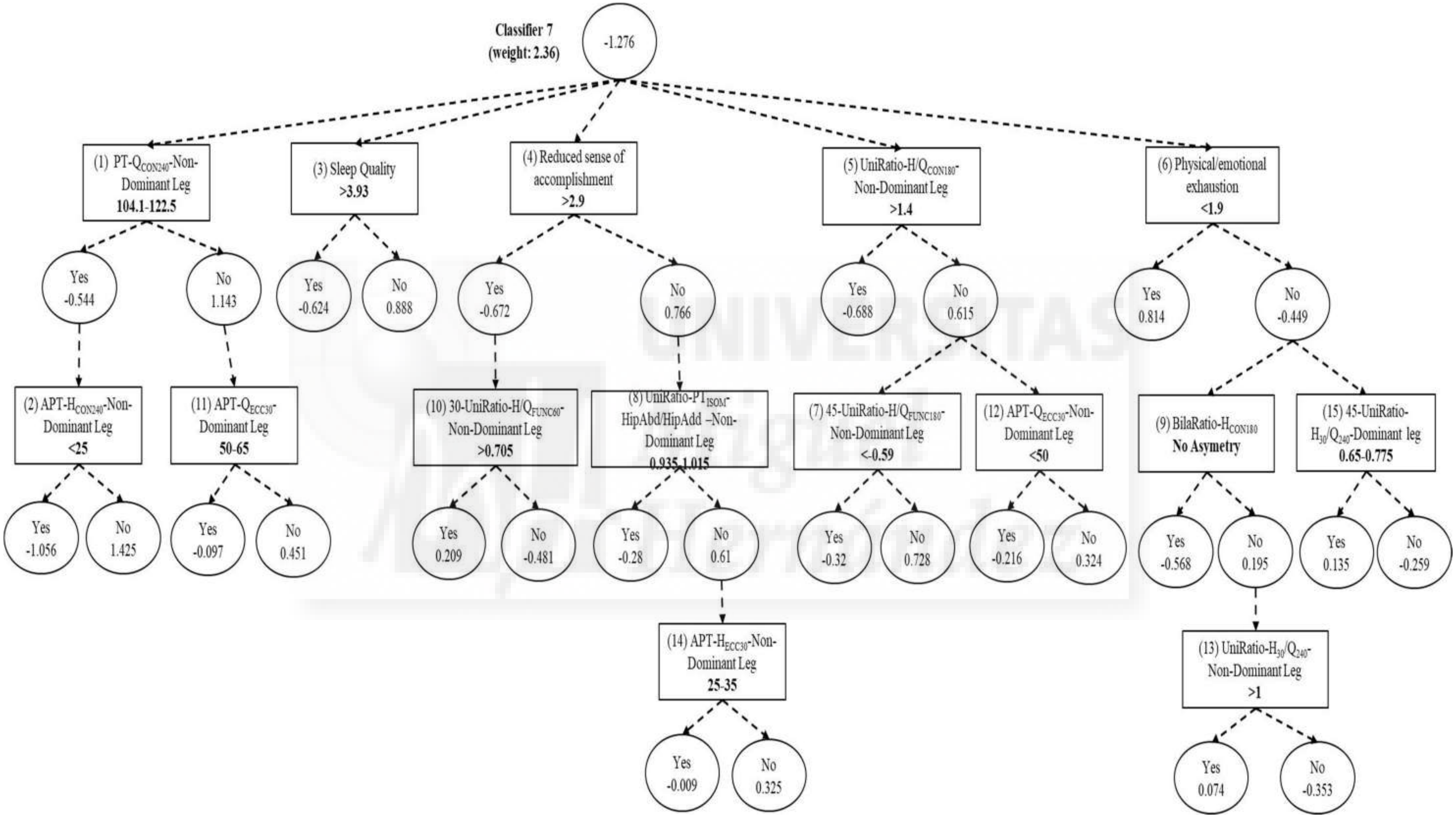


Figure 6.8. Hamstring strain injury predictive model, classifier 7.

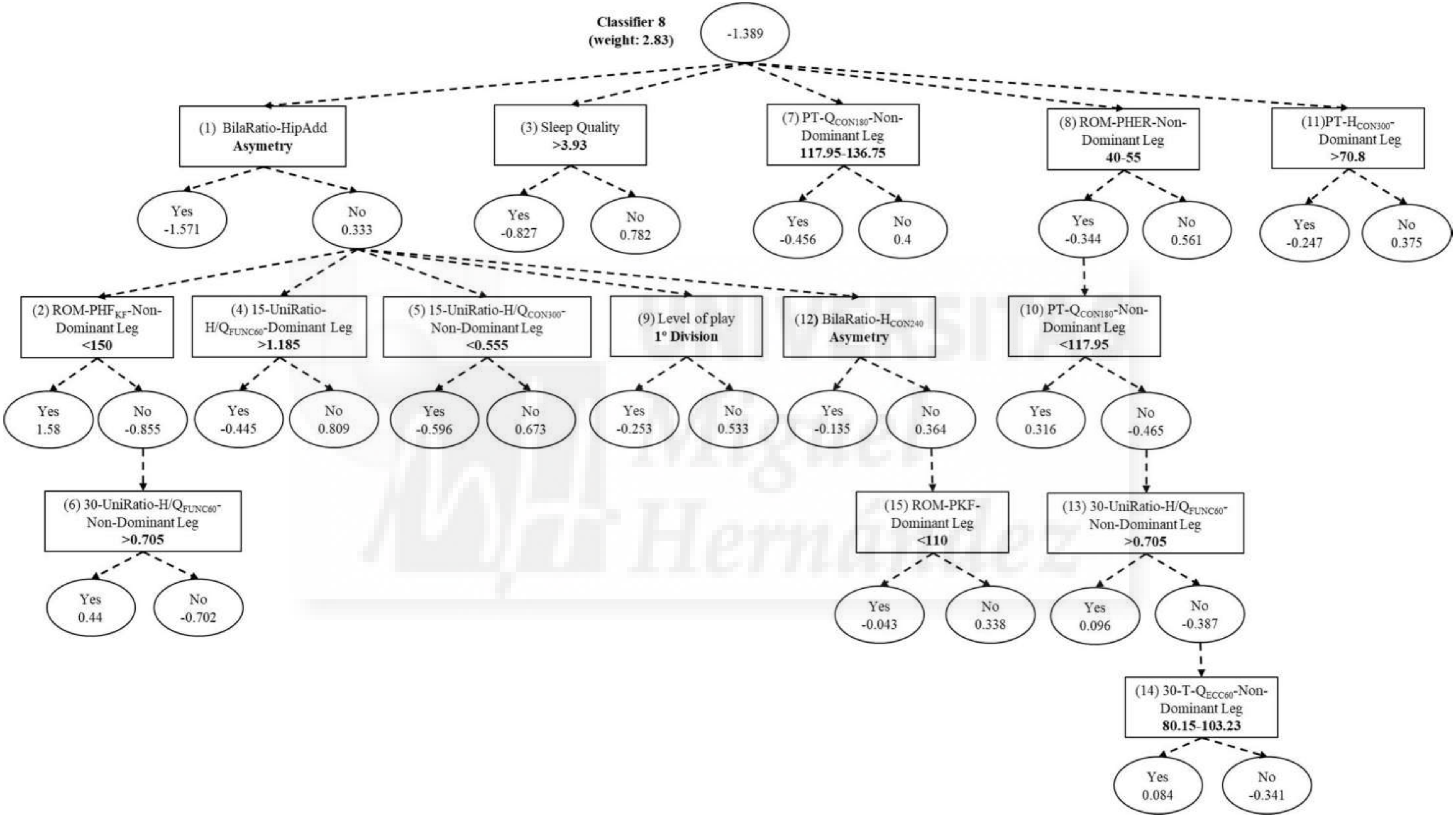


Figure 6.9. Hamstring strain injury predictive model, classifier 8.

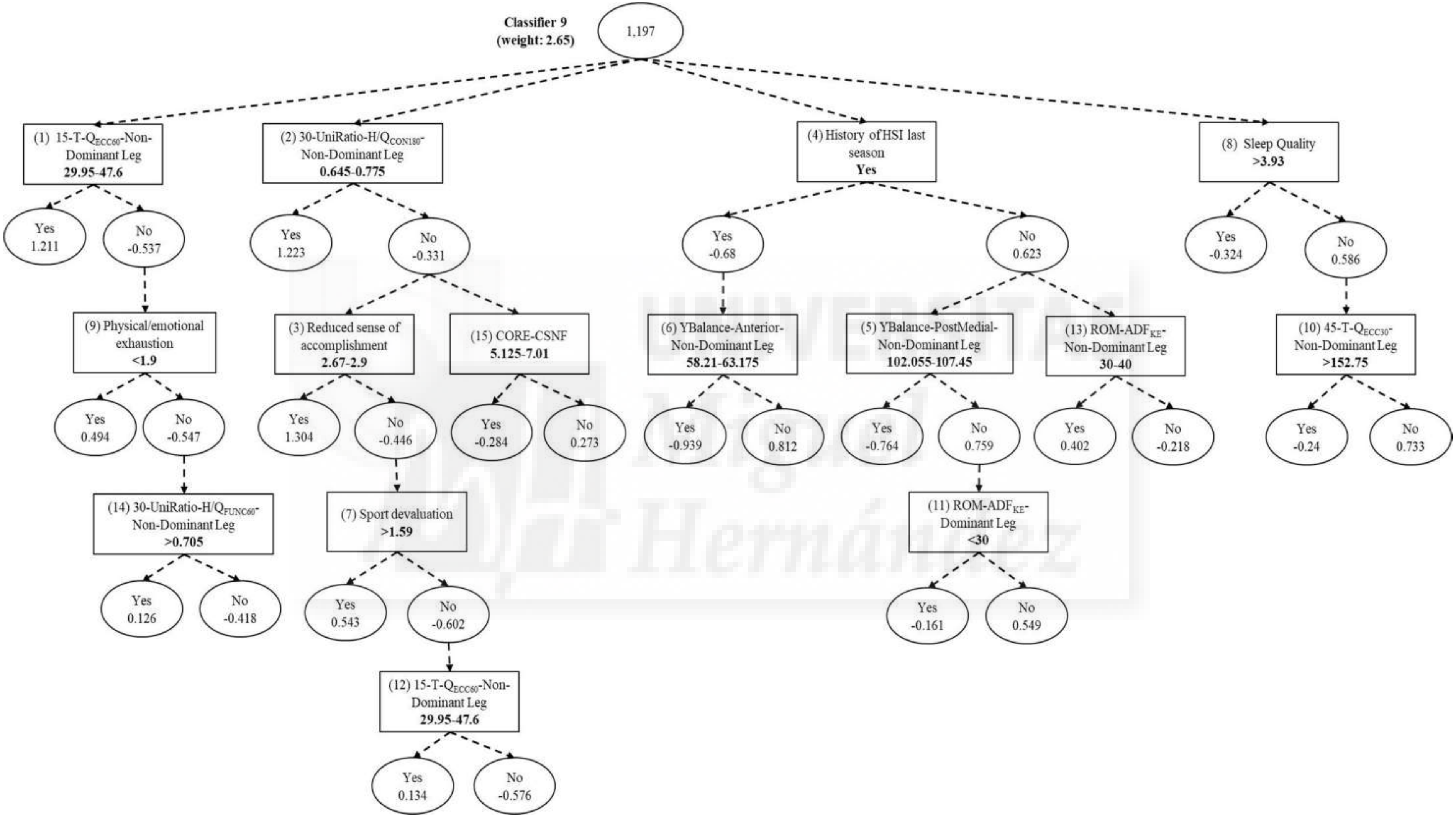


Figure 6.10. Hamstring strain injury predictive model, classifier 9.

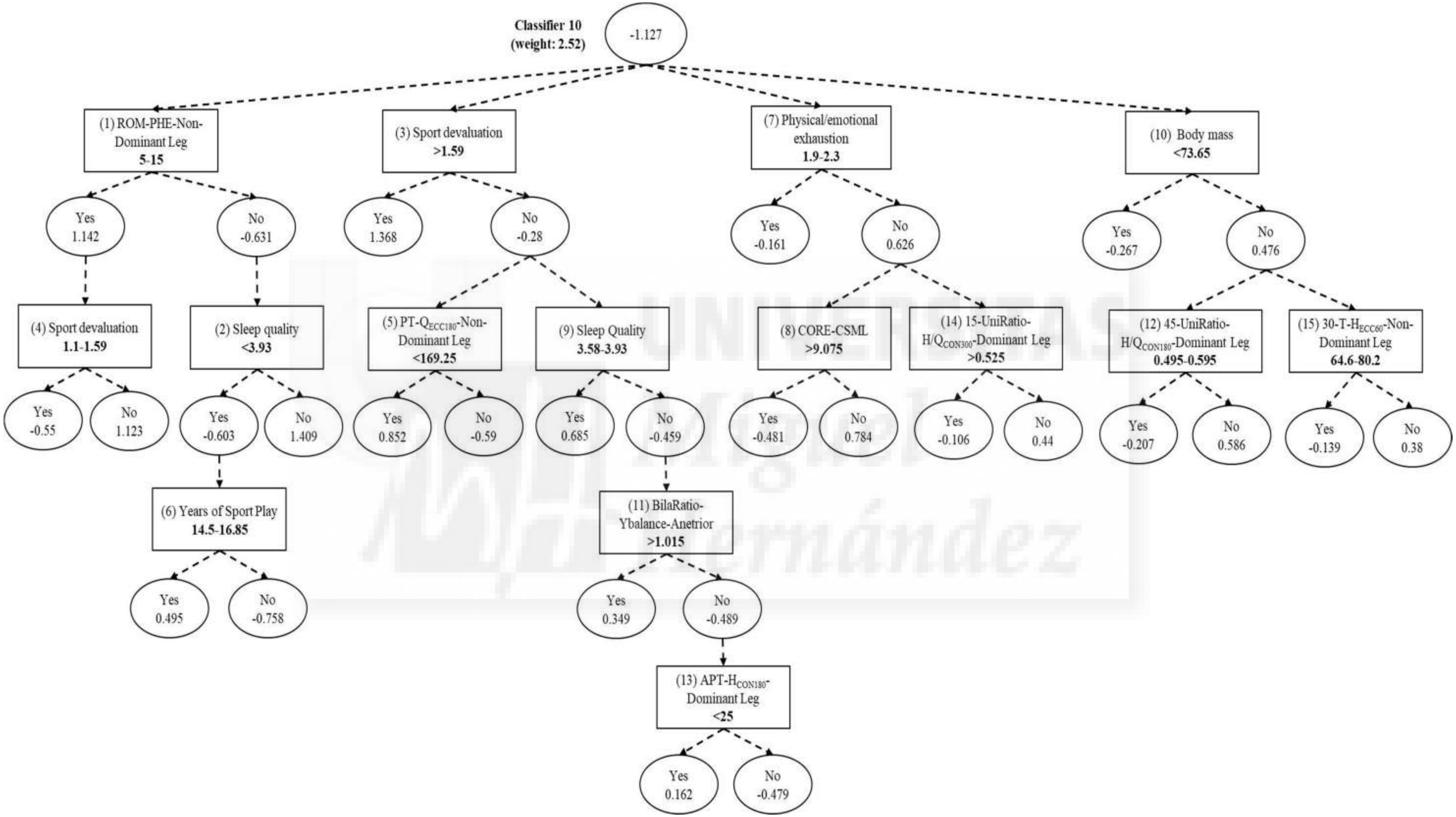


Figure 6.11. Hamstring strain injury predictive model, classifier 10.

The confusion matrix and the main cross validation results of the final model are shown in tables 6.3 and 6.4 respectively.

Table 6.3. Confusion matrix.

A	B	Classified as
15	6	A = Injured
10	89	B = Not Injured

Table 6.4. Cross validation results for the final prediction model.

Correctly classified instances	104 (86.6%)
Incorrectly classified instances	16 (13.3%)
Kappa statistic	0.571
Mean absolute error	0.137
AUC	0.867

AUC: area under the receiver operating characteristic curve.

6.5. Discussion

The current study is the first (to the best of our knowledge) that has used a model to predict HSI by applying a novel multifactorial approach and whose predictive ability has been determined through the exigent resampling technique called cross-validation. In this study the HSI risk model is comprised of 10 classifiers with a tree-shape structure and was developed thanks to the application of learning algorithms (on the training subsets) widely used in the data mining setting. Thus, the model reports an AUC score of 0.867 with false positive and negative rates of 10% and 28% respectively.

The predictive ability of the current model to identify players at high risk of HSI is much higher than those reported in models from previous studies in which less exigent validation processes were applied.[93,118,119,121,122,125,207] Thus, and for example, van Dyk et al.,[125] after having carried out a preseason assessment of the isokinetic hamstring and quadriceps strength in a large cohort of professional soccer players found that in spite of the fact that the regression analysis reported the presence of two independent predictors that were associated with the risk of HSI (hamstring eccentric strength and quadriceps concentric strength), the ROC analysis demonstrated an AUC lower than 0.6. Likewise, Timmins, Bourne, Shield, Williams, Lorenzen, & Opar[207] stated that those soccer players showing eccentric knee flexion strength scores lower than 337 N had 4.4 times greater risk of a subsequent HSI in comparison with stronger players. However, the reported value of the ROC for this cut-off score was only 0.65.

In the current study, the learning process of the model started with 229 features, however the final model only considered 52 of them relevant (Table 6.5). This finding indicates that the range of variables required to identify high and low risk players is manageable in real world settings and would considerably reduce the time required in the preseason screening processes aimed at identifying players at high risk of HSIs. The three main categories of potential injury risk factors employed in the current study (psychological, personal and neuromuscular) all have some representation in the final model selected and hence, this reinforces the idea that the aetiology of HSI is multifactorial.

Table 6.5. Risk factor measures included in the model for predicting hamstring strain injury and the number of times that they appear in the classifiers.

Risk Factor	**Nº of Classifiers**
Personal measures	
Age	1
Body mass	2
History of HSI last season	4
Level of play	1
Sport	1
Years of Sport Play	1
Psychological measures	
Physical/emotional exhaustion	3
Reduced sense of accomplishment	5
Sleep Quality	9
Sport devaluation	3
Dynamic postural control measures	
YBalance-Anterior-Dominant Leg	1
YBalance-Anterior-Non-dominant Leg	1
YBalance-BilaRatio-Anterior	2
YBalance-BilaRatio-PostLateral	1
YBalance-PostLateral-Dominant Leg	1
YBalance-PostLateral-Non-dominant Leg	1
YBalance-PostMedial-Non-dominant Leg	1
Isometric hip abduction and adduction strength measures	
PT$_{ISOM}$- HipAdd -Non-dominant Leg	1
UniRatio-PT$_{ISOM}$-HipAbd/HipAdd –Non-dominant Leg	1
BilaRatio-PT$_{ISOM}$-HipAdd	1
Lower extremity joint range of motion measures	
ROM-ADF$_{KE}$-Dominant Leg	1
ROM-ADF$_{KE}$-Non-dominant Leg	3
ROM-PHA-Dominant Leg	1
ROM- PHE-Non-dominant Leg	2
ROM- PHER-Dominant Leg	1
ROM- PHER-Non-dominant Leg	1
ROM- PHF$_{KE}$-Non-dominant Leg	2
ROM- PHF$_{KF}$-Dominant Leg	1
ROM- PHF$_{KF}$-Non-dominant Leg	1
ROM-PKF-Dominant Leg	1
ROM-PKF-Non-dominant Leg	1

Core stability measures	
CORE-CS$_{ML}$	3
CORE-CS$_{NF}$	3
Isokinetic knee flexion and extension strength measure	
15-UniRatio-H/Q$_{FUNC180}$-Dominant Leg	1
15-UniRatio-H/Q$_{FUNC60}$-Dominant Leg	1
15-UniRatio-H/Q$_{CON60}$-Non-dominant Leg	1
30-T-H$_{ECC180}$-Non-dominant Leg	2
30-T-H$_{ECC60}$-Non-dominant Leg	4
30-UniRatio-H/Q$_{CON180}$-Non-dominant Leg	1
30-UniRatio-H/Q$_{CON240}$-Dominant Leg	1
30-UniRatio-H/Q$_{CON60}$-Dominant Leg	1
30-UniRatio-H/Q$_{FUNC60}$-Non-dominant Leg	4
45-UniRatio-H$_{30}$/Q$_{240}$-Dominant leg	1
45-T-Q$_{ECC30}$- Non-dominant Leg	1
45-T-Q$_{ECC30}$-Dominant Leg	1
45-T-Q$_{ECC60}$-Dominant Leg	1
45-T-H$_{ECC180}$-Non-dominant Leg	1
45-UniRatio-H/Q$_{CON180}$-Dominant Leg	1
45-UniRatio-H/Q$_{CON240}$-Dominant Leg	1
45-UniRatio-H/Q$_{CON300}$-Dominant Leg	1
45-UniRatio-H/Q$_{FUNC60}$-Non-dominant Leg	1
45-UniRatio-H/Q$_{FUNC180}$-Non-dominant Leg	2
APT-Q$_{CON60}$-Dominant Leg	1
APT-H$_{CON180}$-Dominant Leg	1
APT-H$_{CON240}$-Non-dominant Leg	1
APT-H$_{ECC180}$-Dominant Leg	2
APT-Q$_{ECC30}$-Dominant Leg	2
APT-Q$_{ECC30}$-Non-dominant Leg	3
BilaRatio-H$_{CON180}$	1
BilaRatio-H$_{CON240}$	1
PT-H$_{ECC180}$-Non-dominant Leg	1
PT-H$_{ECC60}$-Non-dominant Leg	1
PT-Q$_{ECC180}$-Dominant Leg	1
PT-Q$_{ECC180}$-Non-dominant Leg	1
PT-Q$_{ECC60}$-Dominant Leg	1
UniRatio-H/Q$_{CON180}$-Non-dominant Leg	1
UniRatio-H/Q$_{CON300}$-Non-dominant Leg	2
UniRatio-H/Q$_{CON60}$-Dominant Leg	1
UniRatio-H/Q$_{CON60}$-Non-dominant Leg	2
UniRatio-H/Q$_{FUNC60}$-Dominant Leg	1
UniRatio-H$_{30}$/Q$_{240}$-Dominant Leg	1
UniRatio-H$_{30}$/Q$_{240}$-Non-dominant Leg	1

Abd: abduction; Add: adduction; ADF$_{KE}$: ankle dorsiflexion with knee extended ROM; APT: angle of peak torque; Bila: bilateral; CON: concentric; CS$_{ML}$: unstable sitting while performing medial-lateral displacements with feedback; CS$_{NF}$: unstable sitting without feedback; ECC: eccentric; FUNC: functional; H: hamstring; HSI: hamstring strain injury; ISOM: Isometric; PHA: passive hip abduction ROM; PHE: passive hip extension ROM; PHER: passive hip external rotation ROM; PHF$_{KE}$: passive hip flexion with knee extended ROM; PHF$_{KF}$: passive hip flexion with knee flexed ROM; PKF: passive knee flexion ROM; PT: peak torque; Q: quadriceps; ROM: range of motion; Uni: unilateral.

The main features related to the psychological category of burnout (physical/emotional exhaustion, reduced sense of accomplishment and sport devaluation) were important, but specifically sleep quality was an important risk factor as it was the most consistent variable present in the classifiers (9 out of 10 classifiers). This is the first study that has analysed whether burnout and sleep quality measures are predictive of HSI, alongside other known variables, and therefore direct comparisons are not possible. However, this finding is in concordance with the results found by Cresswell, & Eklund[271] who reported statistically significant correlations between sport-injuries and feelings of sport devaluation in a cohort of professional rugby players. Perhaps, the feeling of frustration experienced by players with a short-term history of HSI might lead them to lose concentration and this can impair the neuromuscular readiness to perform high-intensity intermittent actions during both training and match play, and thus might increase the risk of HSI.

Furthermore, previous HSI, identified by the variable “history of HSI last season” also reported a high presence among the classifiers of the model, evident in four out of ten. This finding is in agreement with the findings of several previous studies,[23,57,93,288] although not all,[20,53] in which previous HSI has been identified as an independent predictor for HSI in professional soccer players. Remaining deficits in physical conditioning or proprioception, or altered movement patterns after a previous injury may provide a plausible link to an anatomically unrelated injury in a following season.[23]

The findings of the current study also highlight the special relevance that the reciprocal hamstring-to-quadriceps ratios, calculated using angle specific torque values close to full extension, present in the identification of players at high risk of HSI in comparison with their homologous ratios calculated by using peak toque values. Likewise, hamstring and quadriceps eccentric torque values obtained close to knee extension (30º and 45º mainly) also seem to adopt a critical role in the predictive model. A possible explanation for this could be attributed to the higher ecological validity of the angle-specific reciprocal H/Q ratios to describe the function of the knee.[292] Biomechanical studies have indicated that hamstring strains are more prone to occur during the latter part of the swing phase of sprinting (closer to full knee extension) when the hamstrings are working eccentrically (energy absorption) to decelerate the knee extension movement (generated among others by the concentric action of the quadriceps muscles) before foot contact, that is, as the muscle develops maximal tension while lengthening to stabilise the knee joint.[244,289] However, peak concentric and eccentric torque production is likely to occur in the mid-late range of the movement (around 40°–80° of knee flexion [0º = full knee extension]).[290] Therefore, this joint angle discrepancy, inherent between any peak torque H/Q ratio and where the HSI is likely to occur, may reduce its validity to assess the muscular balance of the knee. This aspect could justify the reason why the angle-specific H/Q ratios play a more significant role in the likelihood of sustaining a HSI, as they may be more relevant to describe the muscular control of the knee. Although less noticeable, the model built also provides a main role to the isokinetic strength features to predict future HSIs, with 42 features out of 80. These results are not in agreement with the findings showed by van Dyk et al.,[125] and Zvijac, Toriscelli, Merrick, & Kiebzak[291] who did not support the use of isokinetic testing for predicting risk of hamstring injury in subsequent professional competition. Following the same argument wielded for the H/Q ratios, the insufficient ecological validity of the isokinetic methodologies used in the above-mentioned studies could again be a possible reason to explain this discrepancy. Both van Dyk et al.,[125] and Zvijac et al.,[291] examined the relationship isokinetic strength measures and the likelihood of sustaining a hamstring employing isokinetic protocols with the participants adopting a seated position (80°–110° hip flexion). This seated position is not representative of the hip position during sporting tasks (e.g. sprinting, cutting) and does not replicate hamstrings and quadriceps muscle length–tension relationships that occur in the late phase of sprinting, the most hazardous and prone situations to develop a hamstring injury. [244,289] In contrast to these studies, we adopted a prone position (10–20° hip flexion), which has been suggested as being more functionally relevant in term of simulating the injury mechanism.[292]

6.5.1. Clinical implications

In term of practical applications, each classifier has a vote or decision (yes [high risk of HSI] or no [lower risk of HSI]), and the final decision regarding whether or not a player might suffer an injury will be based on the combination of the votes of each individual classifier to each class (yes or no), where the weight of each classifier's vote is a function of its accuracy.

Figures 6.2-6.11 show the weight of the vote of each classifier. Thus, and for example, if a player gets four Yes answers or votes in the classifiers numbers 1, 4, 7 and 9; while the remaining answers to the others classifiers are No, then the final decision will be calculated as follows:

- Yes´ weight = 2.64 (classifier 1) + 3.49 (classifier 4) + 2.36 (classifier 7) + 2.65 (classifier 9) = 11.14
- No´s weight = 2.32 (classifier 2) + 3.2 (classifier 3) + 2.43 (classifier 5) + 2.66 (classifier 6) + 2.83 (classifier 8) + 2.52 (classifier 10) = 15.96
- Final decision = No weight > Yes weight ⇒ YES (high risk of injury)

The ADTree algorithm has the advantage of producing models that are easily represented as a tree with a limited number of nodes (less than 10 in our case). This property is achieved by constructing a tree that is a conjunction of rules which all contribute real-valued evidence towards a given instance being classified as either true (not injured) or false (injured). Unlike traditional tree models the classification of instances by ADTree is thus not determined by a single path traversed in the tree, but rather by the additive score of a collection of paths. The ADTree is graphically represented with two types of nodes: Elliptical *prediction nodes* and rectangular *splitter nodes* (Figure 6.12). Each splitter node is associated with a value indicating the rule condition: If the feature represented by the node satisfied the condition for a given instance, the prediction path will go through the left child node, otherwise the path will go through the right child node. The final classification score produced by the tree is found by summing the values from all the prediction nodes reached by the instance, with the root node being the precondition of the classifier. If the summed score is greater than zero, the instance is classified as true (low risk of injury).

To better explain how coaches and sport practitioners should use the model to predict HSI, we have explained the classifier number 6 or ADTree-6 using the data displayed in table 6.6, which correspond to a fictional soccer player. In addition, figure 6.12 represents in blue the paths followed by the selected instance or example.

Table 6.6. Example data for explaining the model functioning.

Feature	Score
15-UniRatioH/Q_{CON60}-Non-dominant Leg	1.3
30-T-H_{ECC180}-Non-dominant Leg	90.2 Nm
APT-Q_{CON60}-Dominant Leg	52º
45-T-H_{ECC180}-Non-dominant Leg	81 Nm
15-UniRatioH/Q_{CON240}-Non-dominant Leg	0.9
UniRatio-H/Q_{CON60}-Dominant Leg	0.60
History of HSI last season	No
ROM-PHF_{KE}-Non-dominant Leg	75º
UniRatioH/Q_{CON60}-Dominant Leg	0.55
30-UniRatioH/Q_{CON240}-Dominant Leg	0.83
APT-Q_{ECC30}-Dominant Leg	55º
ROM-PHE-Non-dominant Leg	3º

APT: angle of peak torque; CON: concentric; ECC: eccentric; H: hamstring; HSI: hamstring strain injury; Nm: newton per meter; PHE: passive hip extension ROM; PHF_{KE}: passive hip flexion with knee extended ROM; Q: quadriceps; ROM: range of motion; T: torque; Uni: unilateral; º: degrees.

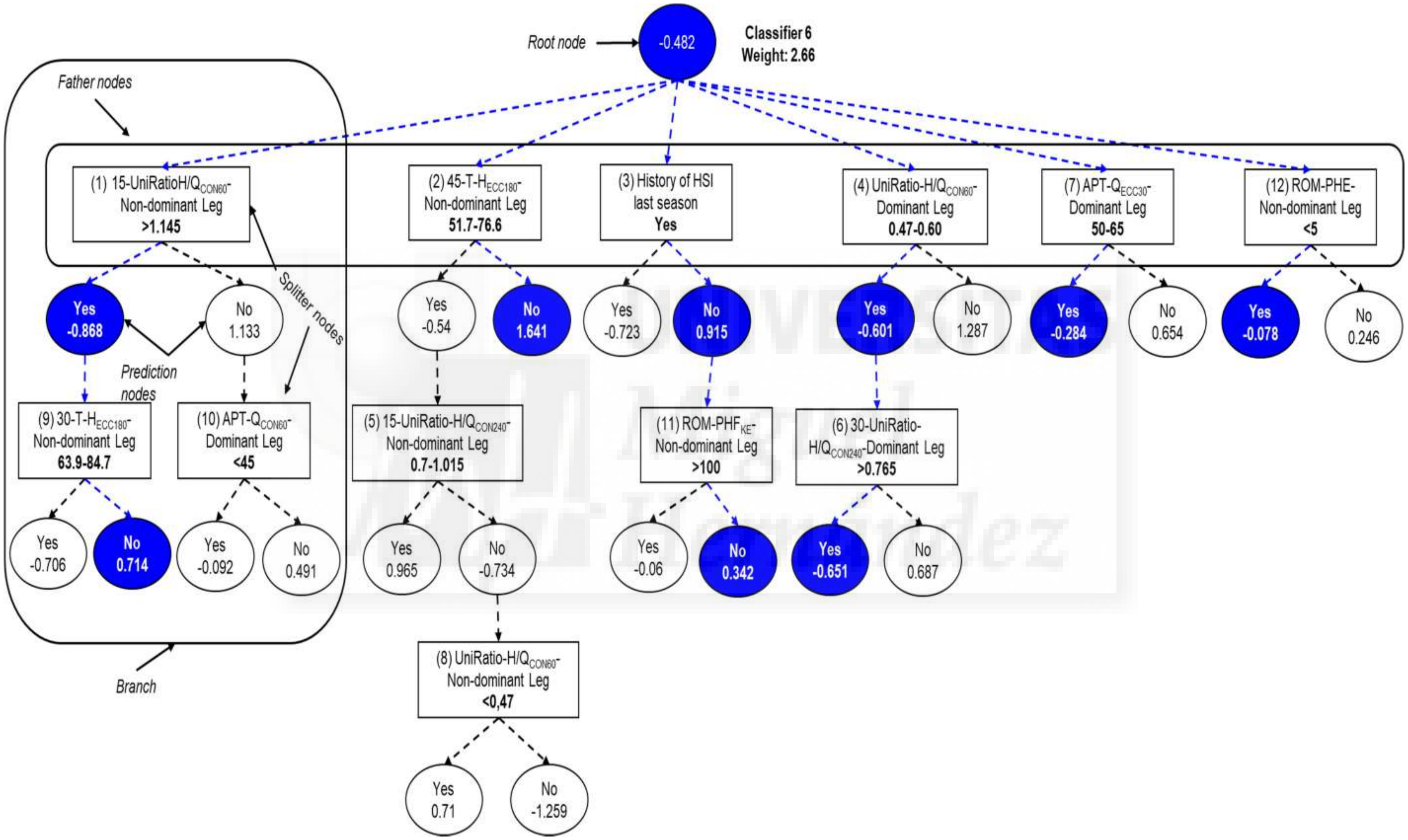


Figure 6.12. Example of alternating decision tree.

This classifier number 6 reports an initial score of -0.482 in its root node. Furthermore, this classifier shows a tree-shape structure comprised by six main branches whose father nodes (first leaves) are the following: a) 15-UniRatioH/Q_{CON60}-Non-dominant Leg; b) 45-T-H_{ECC180}-Non-dominant Leg; c) History of HSI last season; d) UniRatio-H/Q_{CON60}-Dominant Leg; e) APT-Q_{ECC30}-Dominant Leg; and f) ROM-HE-Non-dominant Leg. All the classifiers´ main branches must be addressed, and the scores obtained in each branch (resulting from the data inputted in the father and child [if necessary] nodes) must be summed to the score initially reported by the root node in order to get the final vote of the classifier (yes [high risk of injury] or no [low risk of injury]) for the player.

Thus, and if we start by addressing the branch whose father node is the feature 15-UniRatioH/Q_{CON60}-Non-dominant Leg, it is shown that the score reported by the soccer player (1.3) satisfies the condition present in the node (> 1.145) and hence, he obtains the score of -0.868 from the prediction node Yes. This circumstance drives to the child node represented by the feature 30-T-H_{ECC180}-Non-dominant Leg. In this case, the player does not satisfy the condition presented in the just-mentioned feature, in other words, the value reported (90.2 Nm) is not within the range of 63.9-84.7 Nm. Therefore, here the player achieves a score of 0.714 coming from the predictive node ‘No’. As a consequence, the final result of this branch is the sum of -0.868 plus 0.714, ergo -0.154 points.

The pathway to follow in the branch whose father node is the feature titled 45-T-H_{ECC180}-Non-dominant Leg is shorter than the one previously described, because the player demonstrated a score of 81 Nm, which does not satisfy the established condition (51.7-76.6). Consequently, in this second branch, the player obtains a score of 1.641 from the predictive node ‘No’.The third branch, composed by the father node titled “History of HSI last season” provides a total score of 1.257 (0.915 + 0.342), as the soccer player´s values does not satisfy the condition presented in neither father nor child nodes.

For its part, in the fourth branch, the soccer player does satisfy the condition of the father node, UniRatio-H/Q_{CON60}-Dominant Leg, which provides a score of -0.601. This circumstance drives to the subsequent child node represented by the feature 30-UniRatio-H/Q_{CON240}-Dominant Leg. In this feature, the player obtained a value of 0.83, indicating that satisfy the condition (>0.765), and thus resulting in a score of -0.651. Thus, in this fourth branch, the player achieved a total score of -1.252.

Finally, and for both the fifth and sixth branches, the player again does satisfy the condition presented in their respective father nodes and hence, the scores obtained were 0.284 and -0.078 respectively.

All in all, and after summing the baseline score of the root node with the scores reported in each of the six branches of the classifier, a total score of 0.648 was achieved. This final score is a positive value and this supposes a "No" vote with a weight of 2.66.

6.5.2. Limitations

The model developed in the present study was built with the goal of allowing sport medicine practitioners to accurately identify professional soccer and handball players at high risk of HSI during preseason screenings. To address this issue, we used several predictors (risk factors) as well as external (oversampling) and internal (ensembles) methods and a decision tree (ADTree) as base classifier in order to build a model with very good predictive accuracy. This set up allowed us to build a very powerful model (AUC = 0.867; false positive rate = 10%; false negative rate = 28%) but also very complex in nature (black box approach). Therefore, although the model fulfils the goal for which it was built (to make predictions), its complexity (10 different classifiers and 80 predictors) does not afford the opportunity to answer the question concerning why HSI happens.

Another potential limitation of the current study is the population used. The sport background of participants was professional soccer and handball players and the generalizability to other sport modalities and level of play cannot be ascertained.

Finally, it should also be noted that the model is dependent on the predictors used in the training process and hence, practitioners must follow the same assessment methodologies used in the current study in order to replicate the current results and gain the applicability in their populations.

6.6. Conclusions

To the best of our knowledge this is the first study to use a cross-validation process using data mining techniques to concurrently explore a wide range of HSI risk factors to be able to identify high risk players. This technique appears to permit the identification of high risk players with an AUC value of 0.867, significantly higher than previously reported studies. The current study reinforces that HSI is multifactorial due to the number and range of variables identified in the classifiers. This provides additional challenges for practitioners wanting to screen players and identify them as high or low risk due to the time restraints in real world settings.

6.7. Appendixes

Appendix 6.1. Description of the psychological risk factors recorded.

Name	**Labels**
Sleep quality	< 3.58, 3.58-3.93 or > 3.93
Athlete Burnout Questionnaire	
a) Physical/emotional exhaustion	< 1.9, 1.9-2.3 or ≥ 2.3
b) Reduced sense of accomplishment	< 2.67, 2.67-2.9 or > 2.9
c) Sport devaluation	< 1.1, 1.1-1.59 or > 1.59

Appendix 6.2. Description of the measures obtained from the dynamic postural control test.

Name	**Labels**	
	Dominant Leg	**Non-dominant Leg**
YBalance-Anterior	< 57.46, 57.46-62.555 or > 62.555	< 58.21, 58.21-63.175 or > 63.175
YBalance-Posteromedial	< 101.03, 101.03-107.365 or > 107.365	< 102.055, 102.055-107.45 or > 107.45
YBalance-Posterolateral	< 96.68, 96.68-102.63 or > 102.63	< 95.675, 95.675-101.62 or > 101.62
BilaRatio-YBalance-Anterior	<0.965, 0.965-1.015 or 1.015	
BilaRatio-YBalance-Posteromedial	<0.975, 0.975-1.015 or > 1.015	
BilaRatio-YBalance-Posterolateral	<0.985, 0.985-1.035 or > 1.035	
YBalance-Composite	< 85.4, 85.4-90.075 or > 90.075	< 85.87, 85.87-90.39 or > 90.39

Bila: bilateral

Appendix 6.3. Description of the measures obtained from the isometric hip abduction and adduction strength test.

Name	Labels	
	Dominant Leg	**Non-dominant Leg**
PT_{ISOM}-HipAbd	< 190.05, 190.05-217.625 or > 217.625	< 197, 197-220.585 or > 220.585
PT_{ISOM}-HipAbd-Normalised	< 2.505, 2.505-2.86 or > 2.86	< 2.575, 2.575-2.875 or > 2.875
PT_{ISOM}-HipAdd	< 193.225, 193.225-219.825 or > 219.825	< 189.715, 189.715-219.3 or > 219.3
PT_{ISOM}-HipAdd-Normalised	< 2.59, 2.59-2.92 or > 2.92	< 2.525, 2.525-2.885 or > 2.885
UniRatio-ISOM-HipAbd/HipAdd	No Asymmetry (< 10%) or Asymmetry (≥ 10%)	

Abd: abduction; Add: adduction; Bila: bilateral; ISOM: isometric; PT: peak torque; Uni: unilateral.

Appendix 6.4. Description of the measures obtained from the lower extremity range of motion assessment tests.

Name	Labels
PHF_{KF}	≤ 150 or > 150 (1)
PHF_{KE}	< 80, 80-100 or > 100 (2)
PHE	< 5, 5.0-15 or > 15 (5)
PHA	< 50, 50-70 or > 70 (3)
PHIR	< 45, 45-60 or > 60 (1)
PHER	< 40, 40-55 or > 55 (1)
PKF	< 110, 110-130 or > 130
ADF_{KE}	< 30, 30-40 or > 40 (5)
ADF_{KF}	< 30, 30-40 or > 40 (4)

ROM: range of motion; PHF_{KF}: passive hip flexion with knee flexed ROM; PHF_{KE}: passive hip flexion with knee extended ROM; PHE: passive hip extension ROM; PHA: passive hip abduction ROM; PHIR: passive hip internal rotation ROM; PHER: passive hip external rotation ROM; PKF: passive knee flexion ROM; ADF_{KE}: ankle dorsiflexion with knee extended ROM; ADF_{KF}: ankle dorsiflexion with knee flexed ROM.

(1): American Academy of Orthopaedic Association, 1965[225]; (2): Palmer & Epler, 2002[272]; (3): Gerhardt, 2002[237]; (4) Pope, Herbert & Kirwan (1998)[234]; (5) Cejudo, 2016.[224]

Appendix 6.5. Description of the measures obtained from the core stability test.

Name	**Labels**
CS_{NF}	< 5.125, 5.125-7.01 or > 7.01
CS_{WF}	< 4.795, 4.795-6.04 or > 6.04
CS_{ML}	< 7.555, 7.555-9.075 or > 9.075
CS_{AP}	< 7.505, 7.505-9.075 or > 9.075
CS_{CD}	< 9.47, .47-11.365 or > 11.365
GLOBAL	< 7.15, 7.15-8.8 or > 8.8

CS_{NF}: unstable sitting without feedback; CS_{WF}: unstable sitting with feedback; CS_{ML}: unstable sitting while performing medial-lateral displacements with feedback; CS_{AP}: unstable sitting while performing anterior-posterior displacements with feedback; CS_{CD}: unstable sitting while performing circular displacements with feedback.

Appendix 6.6. Description of the measures obtained from the isokinetic hamstring and quadriceps strength assessment.

Measure	Labels	
	Dominant Leg	**Non-dominant Leg**
	Concentric Muscle Actions	
PT-Q$_{60}$	< 174.4, 174.4-196.8 or > 196.8	< 164.25, 164.25-193.05 or > 193.05
PT-H$_{60}$	< 79.45, 9.45-96.15 or > 96.15	< 73.65, 73.65-90.45 or > 90.45
PT-Q$_{180}$	< 119.3, 119.3-137.55 or > 137.55	< 117.95, 117.95-136.75 or > 136.75
PT-H$_{180}$	< 64.6, 64.6-77.85 or > 77.85	< 63.75, 63.75-76.25 or > 76.25
PT-Q$_{240}$	< 107.15, 107.15-126.3 or > 126.3	< 104.1, 104.1-122.5 or > 122.5
PT-H$_{240}$	< 61.55, 61.55-74.075 or > 74.075	< 59.2, 59.2-70.95 or > 70.95
PT-Q$_{300}$	< 97.9, 97.9-112.15 or > 112.15	< 96.25, 96.25-110.65 or > 110.65
PT-H$_{300}$	< 57.5, 57.5-70.8 or > 70.8	< 52.15, 52.15-63.1 or > 63.1
APT-Q	< 45, 45-60 or > 60	
APT-H	< 25, 25-35 or > 35	
	Eccentric Muscle Actions	
PT-H$_{30}$	< 80.35, 80.35-101.25 or > 101.25	< 75.4, 75.4-89.2 or > 89.2
PT-Q$_{30}$	< 185.45, 185.45-228.35 or > 228.35	< 170.2, 170.2-221.45 or > 221.45
15-T-H$_{30}$	< 62.45, 62.45-89.8 or > 89.8	< 61.2, 61.2-80.65 or > 80.65
15-T-Q$_{30}$	< 28.15, 28.15-44 or > 44	< 30.15, 30.15-46.15 or > 46.15
30-T-H$_{30}$	< 67.55, 67.55-85.85 or > 85.85	< 63.6, 63.6-77.9 or > 77.9
30-T-Q$_{30}$	< 83.75, 83.75-108.05 or > 108.05	< 83.5, 83.5-101 or > 101
45-T-H$_{30}$	< 63.35, 63.35-78.65 or > 78.65	< 58.45, 58.45-70.55 or > 70.55
45-T-Q$_{30}$	< 133.95, 133.95-160.9 or > 160.9	< 124.05, 124.05-152.75 or > 152.75
PT-H$_{60}$	< 81.2, 81.2-101.05 or > 101.05	< 76.55, 76.55-91.45 or > 91.45
PT-Q$_{60}$	< 191.9, 191.9-230.35 or > 230.35	< 181, 181-220.75 or > 220.75
15-T-H$_{60}$	< 68.6, 68.6-87.95 or > 87.95	< 62.3, 62.3-83.05 or > 83.05
15-T-Q$_{60}$	< 30.9, 30.9-42.55 or > 42.55	< 29.95, 29.95-47.6 or > 47.6
30-T-H$_{60}$	< 70.2, 70.2-88.9 or > 88.9	< 64.6, 64.6-80.2 or > 80.2
30-T-Q$_{60}$	< 82, 82-101.45 or > 101.45	< 80.15, 80.15-103.23 or > 103.23
45-T-H$_{60}$	< 66.1, 66.1-83 or > 83	< 62.6, 62.6-77.25 or > 77.25
45-T-Q$_{60}$	< 131.6, 131.6-157.05 or > 157.05	< 130, 130-155.5 or > 155.5
PT-H$_{180}$	< 79.35, 79.35-98.7 or > 98.7	< 75.55, 75.55-90.55 or 90.55
PT-Q$_{180}$	< 177.55, 177.55-211.8 or > 211.8	< 169.25, 169.25-202.25 or > 202.25
15-T-H$_{180}$	< 50.95, 50.95-74.05 or > 74.05	< 51.75, 51.75-76.6 or > 76.6
15-T-Q$_{180}$	< 40.55, 40.55-50.7 or > 50.7	< 41.65, 41.65-55.45 or > 55.45
30-T-H$_{180}$	< 69.15, 69.15-85.95 or > 85.95	< 63.9, 63.9-84.7 or > 84.7
30-T-Q$_{180}$	< 97.1, 97.1-115 or > 115	< 92.35, 92.35-116.2 or > 116.2
45-T-H$_{180}$	< 75.45, 75.45-89.15 or > 89.15	< 71.6, 71.6-84.55 or > 84.55
45-T-Q$_{180}$	< 149.9, 149.9-170.55 or > 170.55	< 142.55, 142.55-171.53 or > 171.53
APT-H	< 25, 25-35 or > 35	
APT-Q	< 50, 50-65 or > 65	
	Unilateral Conventional Ratios	
H/Q$_{CONV60}$	<0.47, 0.47-0.60 or >0.60	
H/Q$_{CONV180}$	≤ 0.60 or > 0.60	
H/Q$_{CONV240}$	≤ 0.60 or > 0.60	
H/Q$_{CONV300}$	<0.6 0.6-0.8 or >0.8	

	Angle-Specific Unilateral Conventional Ratios	
15-H/Q_{CONV60}	<0.935, 0.935-1.155 or > 1.155	<0.915, 0.915-1.145 or > 1.145
15-H/$Q_{CONV180}$	< 1.045, 1.045-1.395 or > 1.395	< 1.045, 1.045-1.4 or > 1.4
15-H/$Q_{CONV240}$	<0.77, 0.77-1.15 or > 1.15	<0.7, 0.7-1.015 or > 1.015
15-H/$Q_{CONV300}$	<0.525, 0.525-0.87 or >0.87	<0.555, 0.555-0.83 or >0.83
30-H/Q_{CONV60}	<0.635, 0.635-0.745 or >0.745	<0.625, 0.625-0.725 or >0.725
30-H/$Q_{CONV180}$	<0.685, 0.685-0.82 or >0.82	<0.645, 0.645-0.775 or >0.775
30-H/$Q_{CONV240}$	<0.645, 0.645-0.765 or >0.765	<0.635, 0.635-0.745 or >0.745
30-H/$Q_{CONV300}$	<0.845, 0.845-1.085 or > 1.085	<0.845, 0.845-1.045 or > 1.045
45-H/Q_{CONV60}	<0.445, 0.445-0.515 or >0.515	<0.435, 0.435-0.515 or >0.515
45-H/$Q_{CONV180}$	<0.495, 0.495-0.595 or >0.595	<0.485, 0.485-0.555 or >0.555
45-H/$Q_{CONV240}$	<0.525, 0.525-0.625 or >0.625	<0.505, 0.505-0.585 or >0.585
45-H/$Q_{CONV300}$	<0.555, 0.555-0.635 or >0.635	<0.495, 0.495-0.595 or >0.595
	Unilateral Functional Ratios	
H/Q_{FUNC60}	<0.6, 0.6-0.7 or >0.7	
H/$Q_{FUNC180}$	≤ 0.80 or > 0.80	
H_{30}/Q_{240}	<0.8, 0.8-1.0 or > 1.0	
	Angle-Specific Unilateral Functional Ratios	
15-H/Q_{FUNC60}	<0.905, 0.905-1.185 or > 1.185	<0.895, 0.895-1.145 or > 1.145
15-H/$Q_{FUNC180}$	<0.875, 0.875-1.315 or > 1.315	<0.965, 0.965-1.355 or > 1.355
15-H_{30}/Q_{240}	< 1.395, 1.395-1.83 or > 1.83	< 1.22, 1.22-1.655 or > 1.655
30-H/Q_{FUNC60}	<0.615, 0.615-0.745 or >0.745	<0.585, 0.585-0.705 or >0.705
30-H/$Q_{FUNC180}$	<0.755, 0.755-0.945 or >0.945	<0.725, 0.725-0.875 or >0.875
30-H_{30}/Q_{240}	<0.875, 0.875-1.095 or > 1.095	<0.805, 0.805-0.965 or >0.965
45-H/Q_{FUNC60}	<0.445, 0.445-0.535 or >0.535	<0.425, 0.425-0.495 or >0.495
45-H/$Q_{FUNC180}$	<0.665, 0.665-0.785 or >0.785	<0.59, 0.59-0.705 or >0.705
45-H_{30}/Q_{240}	<0.65, 0.65-0.775 or >0.775	<0.595, 0.595-0.71 or >0.71
	Bilateral Ratios	
H/H_{CON60}	No Asymmetry or Asymmetry	
H/H_{CON180}	No Asymmetry or Asymmetry	
H/H_{CON240}	No Asymmetry or Asymmetry	
Q/Q_{CON60}	No Asymmetry or Asymmetry	
Q/Q_{CON180}	No Asymmetry or Asymmetry	
Q/Q_{CON240}	No Asymmetry or Asymmetry	
H/H_{ECC60}	No Asymmetry or Asymmetry	
H/H_{ECC180}	No Asymmetry or Asymmetry	

APT: angle of peak torque; CON: concentric; CONV: conventional; FUNC: functional; ECC: eccentric; KE: knee extension; KF: knee flexion; PT: peak torque.

Appendix 6.7.

Data pre-processing

To optimise the performance of the different learning algorithms used in the data processing stage, standard pre-processing methods such as data cleaning and data discretization were applied.

Firstly, those players who did not complete all the neuromuscular tests for any reason (six soccer players) were removed. Furthermore, four soccer players were also removed because they left their respective teams before the follow up procedure was completed. Secondly, an investigation regarding the presence of outliers was carried out using boxplots and the detected outliers were removed. The third step consisted in looking for missing data. To address this issue, frequency tables and diagrams were built. Thus, missing data were replaced by the mean value of the corresponding feature of the specific sport modality (soccer or handball) of the players. For example, if a soccer player did not report his height for any reason, then the average value of his counterpart soccer players was inputted. It should be pointed out that none of the features reported a percentage of missing data and outliers higher than 5%. The SPSS Statistical software (V21.0) was used to carry out these data cleaning processes.

After having applied the above-mentioned data cleaning methods, an imbalance data set (showing an imbalance ratio of 0.34) comprised of 88 soccer and 34 handball players (instances) and 229 potential risk factors (features) was created.

The final step comprised the discretization of the continuous features as this has been shown to be an effective measure to improve the performance of several classifiers.[264,293] Thus, continuous features were discretized applying the unsupervised discretization algorithm available in the well-known Weka (Waikato Environment for Knowledge Analysis) data mining software[294] and using the equal frequency binning approach (three intervals). We selected three intervals in order to reflect taxonomy of low, moderate and high scores that might make the final models more comprehensible. In those features where the graphical representation of the data allows the authors to suggest alternative cut-off values, a comparative analysis was run in order to identify the discretization approaches (algorithm vs. authors' visual inspection) that displayed the best predictive ability. The approach reporting the better predictive results was used for the discretization of each feature. Consequently, lower extremity ROM and isokinetic APT features as well as both the reciprocal knee flexion to knee extension ratios and bilateral knee flexion and extension ratios were discretized using the graphical representation of the data as a guide; whereas the remaining features were discretized using the Weka unsupervised discretization algorithm (appendixes 6.1-6.6).

Data processing

Part of the taxonomies for external (oversampling) and internal (ensembles) methods for learning with imbalanced data sets proposed by Galar et al.,[113] López et al.,[114] and Elkarami et al.,[263] was used to build models for predicting HSI in professional soccer and handball players. Thereby, the algorithms of each of the above mentioned families (oversampling and ensembles) that showed the best goodness scores in the latter mentioned studies were used to train models. The model with the highest validity metrics was considered the best for predicting HSI based on the current data set.

To achieve founded conclusions, three decision tree algorithms were selected to be used in the oversampling and ensemble methodologies as base classifiers: J48,[265] which is an algorithm for generating a pruned or unpruned C4.5 decision tree; ADTree,[267] which is an alternating decision tree; and SimpleCart,[266] which implements minimal cost-complexity pruning.

All the decision trees selected were made cost sensitive to minimize the cost of misclassification of the minority class by using the filter cost sensitive classifier algorithm available in Weka workbench. Thus, the training data were reweighted according to the costs assigned to each class. The set up of the definitive cox matrix was based on the best performance reported after testing all the possibilities. For the sake of brevity and the lack of space, the codes of the algorithms used in this study are not presented. Instead, only the names of the algorithms have been specified and the reader is referred to the original sources. Furthermore, all the classification algorithms used are available in the Weka software.[294]

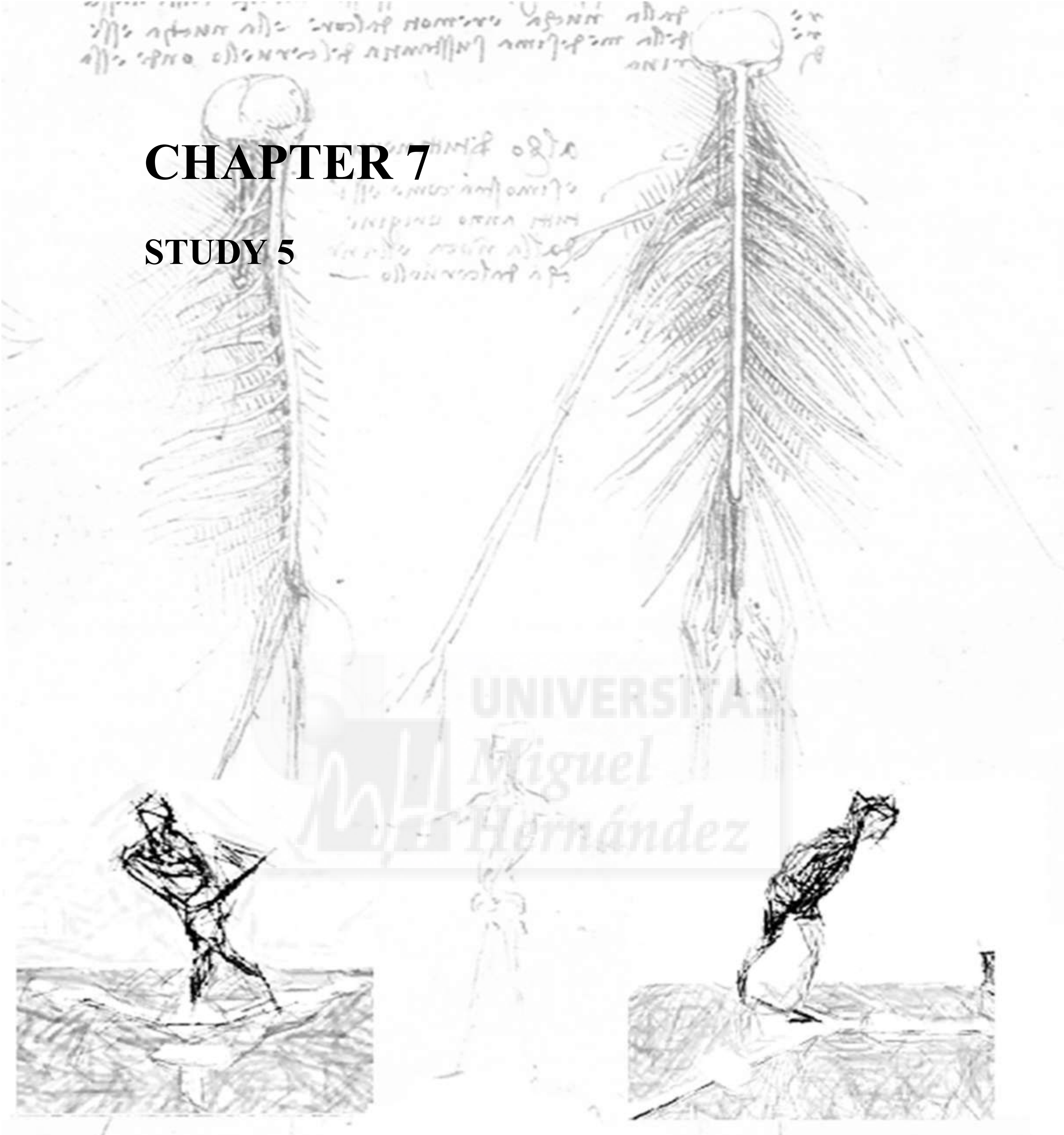
Although there are several data oversampling methods, we used one of the most popular methodologies that is the classic synthetic minority oversampling technique (SMOTE).[295] The main concept behind SMOTE is to create new minority class examples by interpolating several minority class instances that lie together for oversampling the training set. With this technique, the positive class is oversampled by taking each minority class sample and introducing synthetic examples along the line segments joining any/all of the k minority class nearest neighbours. Three different levels of balance in the training data were analysed (25:75; 40:60; 50:50) and the best in term of predictive ability was reported. Additionally, the interpolations that are computed to generate new synthetic data were made considering the 5-nearest neighbours of minority class instances using the Euclidean distance.

Regarding ensemble learning algorithms, the algorithm families designed to deal with skewed class distributions in data sets were included: Boosting-based and Bagging-based. The Boosting-based ensembles that were considered in the current study were SMOTEBoostM1[296] and RUSBoost.[297] With respect to Bagging-based ensembles, it was included from the

OverBagging group, OverBagging (which uses random oversampling) and SMOTEBagging.[298] In order to evaluate the performance of the decision tree algorithms, the worldwide-accepted fivefold SCV technique was used. That is, we split the dataset into fivefolds, each one containing 20% of the patterns of the dataset. For each fold, the algorithm was trained with the examples contained in the remaining folds and then tested with the current fold. A wide range of classification performance measures can be obtained from the SCV technique. A well-known approach to unify these measures and to produce an evaluation criterion is to use the area under the ROM curve (AUC). In particular, the AUC corresponds to the probability of correctly identifying which one of the two stimuli is noise and which one is signal plus noise.[114] Thus, the AUC was used as a single measure of a classifier's performance for evaluating which model is better on average. Furthermore, two extra measures from the confusion matrix were also used as evaluation criteria: a) true positive rate (TPrate): TPrate = $\frac{TP}{TP + FN}$ also called sensitivity or recall, is the proportion of actual positives which are predicted to be positive; and b) true negative rate (TNrate): TNrate = $\frac{TN}{TN + FP}$ or specificity, is the proportion of actual negatives which are predicted to be negative.

CHAPTER 7

STUDY 5



Under review in Scandinavian Journal of Medicine & Science in Sports

Relationships and sex-related differences in several neuromuscular parameters with dynamic balance in soccer players

Alejandro Lopez-Valenciano, Francisco Ayala Rodríguez, Mark De Ste Croix, David Barbado Murillo, Francisco J. Vera-Garcia

CHAPTER 7

STUDY 5

Relationships and sex-related differences in several neuromuscular parameters with dynamic balance in soccer players

by

Lopez-Valenciano A, Ayala F, De Ste Croix M, Barbado D, & Vera-Garcia FJ

7.1. Abstract

Background: Unilateral dynamic balance (UDB) has been suggested to be critical in performing several explosive sport actions, as well as to identify players at high risk of non-contact lower extremity injury. To design training interventions to improve dynamic balance, knowledge of the relationships between dynamic balance and specific neuromuscular factors such as knee and hip strength, lower extremity joint range of motion (ROM) and core stability is essential.

Purpose: The purpose of this study was to analyse the relationships between several parameters of neuromuscular performance with UDB measured throughout Y-Balance test (Y-BT), as well as to determine the possible sex-related differences in a cohort of professional soccer players.

Methods: The Y-BT, isokinetic (concentric and eccentric), knee flexion and extension strength, isometric hip abduction and adduction strength, lower extremity joint ROMs (hip, knee and ankle) and trunk stability were assessed in male (n = 88) and female (n = 44) professional soccer players. A stepwise multivariate linear least square regression with backward elimination analysis was carried out to identify a group of factors that were independently associated with balance performance in both sexes.

Results: The results showed that passive hip flexion and ankle dorsiflexion with knee flexed (ADF_{KF}) ROM were the main factors that retained a significant association to dominant (R^2 = 23.1) and non-dominant (R^2 = 33.5) composite scores for males. For females, trunk stability (unstable sitting while performing medial-lateral displacements with feedback), hip abduction isometric peak torque and passive hip abduction and ADF_{KF} ROM variables retained a significant association for both, dominant (R^2 = 38.2) and non-dominant (R^2 = 46.9) legs.

Conclusions: training interventions aimed at improving or maintaining UDB in professional male soccer players should include, among others, stretching exercises for the posterior chain of the lower extremity. However, females should also include exercises for strength and mobility of the hip abductors and core stability (especially in the frontal plane).

Keywords: *Y-Balance, injury, strength, core stability, performance.*

7.2. Introduction

Unilateral dynamic balance (UDB), defined as the ability of an individual to maintain the center of mass within the body´s base whilst performing single leg movements,[140] is considered a fundamental ability in the organization of skilled motor performance.[299] It has been demonstrated that good UDB is an essential prerequisite to safely and accurately perform several explosive sport actions carried out over a single leg, such as sudden acceleration and deceleration tasks, rapid changes of direction, kicking and jumping.[300]

Although several sophisticated assessment methodologies have been proposed to measure UDB (e.g. force platforms, 3D motion analysis devices), the field based Y-Balance test (Y-BT) appears to be the most popular in clinical, research and sport settings for several reasons. For instance, it has been considered operationally valid because it offers sufficient challenge for UDB as the subject must maintain balance on a single leg, whilst the other leg carries out a series of reaching tasks.[301] In addition, the Y-BT has been shown to be sensitive enough: a) to detect UDB deficits in patients with chronic ankle instability,[302] patellofemoral pain syndrome[303] and anterior cruciate ligament (ACL) deficiency[304]; b) to identify athletes at high risk of non-contact lower extremity injury[97,98]; and c) to monitor the rehabilitation and return to play processes.[305] Furthermore, the test has also been shown high intra and inter-tester reliability.[306] Finally, the Y-BT may be considered a clinically efficient, field-ready test because its procedure is simple to administer, instructions are easy to follow, scores are easy to explain, the movements require minimal skills training and large numbers can be tested in a short period of time.[301]

Therefore, due to the relevance of the Y-BT (as a measure of UDB) for sport performance and injury prevention and rehabilitation, it seems necessary to identify which measures of neuromuscular performance (e.g. hip and knee strength, lower extremity joint range of motion [ROM], core stability [CS]) could have an impact on its scores in order to design targeted training interventions.

Furthermore, the study of the relationships between the measures of neuromuscular performance with UDB should be specific to each sport modality and competition level due to the differences in technical skills, specific movements, training load and physical capacities between sports. These sport specific adaptations through training and competition may predispose participants to individual chronic musculoskeletal adaptations, thus possibly developing different strategies for neuromuscular control and influencing subsequent Y-BT scores.[97] In this sense, elite soccer players demonstrated better UDB capability than their non-elite peers[143,307,308] and when compared with other sporting populations,[309] suggesting that the Y-BT may be sensitive to training status and/or sport-related adaptations.

Professional soccer players might be one of the target populations of UDB training programmes since they are required to perform repetitively high intensity explosive actions, many of them executed from a single leg (e.g. kicking, jumping and landing tasks, rapid changes of direction)[308,309] that place considerable demands on UDB. Therefore, the stability of the stance foot in the execution of successful soccer related movement might be crucial.[300] Furthermore, epidemiology studies have shown that professional soccer players present one of the highest reported incidence rates of non-contact ankle and knee injuries, where a poor UDB has been suggested as a primary risk factor.[310]

Although some studies have explored the individual contribution of certain modifiable measures of neuromuscular performance on Y-BT in soccer (knee[141-143] and hip[144] strength, jumping ability,[141,145] core stability,[145] ankle dorsiflexion[146,147] and hip flexion[147] ROMs) only one study has used professional players.[141] In addition, to the authors´ knowledge, no studies have analysed the concurrent influence of the main training modifiable neuromuscular measures (hip and knee strength, CS and lower extremity joint ROM) in the Y-BT performance in soccer players.

Finally, and although previous studies have reported no sex-related differences in Y-BT reached distances in college athletes,[148] basketball players[149] and recreational athletes,[150] none of them have determined if males and females used similar or different neuromuscular strategies to achieve them. Consequently, the relationships between the main training modifiable measures of neuromuscular performance with Y-BT in professional soccer players remain unresolved. This knowledge would allow clinicians and sport practitioners to develop more effective and tailored UDB training programmes in soccer players, possibly improving performance and reducing the risk of injury.

Therefore, the main purpose of this study was to analyse the relationships between several parameters of neuromuscular performance with UDB measured throughout Y-BT, as well as to determine the possible sex-related differences in a cohort of professional soccer players.

7.3. Methods

7.3.1. Sample size estimation

The sampling software package GPower 3.1 (sample size estimation, contrast of hypothesis, comparing groups' means, independent groups) was used to calculate (a priori) the sample size needed to detect meaningful results through a linear multiple regression analysis. An alpha level of 0.05, a desired power of 0.9, an effect size of 0.02 (weak) and 5 predictors were introduced as input in the sample size estimation analysis.

The analysis indicated that a minimal sample size of 45 participants would be required for each group (males and females). Considering the possible level of dropout in this type of intervention (around 25%) and that typically 18-22 players comprise a typical professional soccer team, players recruited from 4 different teams for each group would be needed to ensure an appropriate final sample size.

7.3.2. Participants

A total of 88 male and 79 female professional soccer players were contacted to take part in the current study (convenience sampling). To be included, all participants had to be free of pain at the time of the study and currently involved in soccer-related activities. Participants were excluded if they reported the presence of any lower extremity injury within the last month, a current upper respiratory tract infection, any bone or joint abnormalities, any uncorrected visual and vestibular problems and/or a concussion within the last three months.[141] The study was conducted during the pre-competitive phase of the year 2013.

Before any participation, experimental procedures and potential risks were fully explained to the participants in verbal and written form, and written informed consent was obtained from participants. An institutional research ethics committee approved the study protocol prior data collection, conforming to the recommendations of the Declaration of Helsinki.

Of the 79 female players contacted, all female players from two teams (n = 35) were excluded from the study because they did not complete the testing sessions due to time restrictions (one team) and technical problems (one team). Therefore 132 professional soccer players (88 male and 44 female) from 6 different soccer teams completed this study (Table 7.1).

Table 7.1. Demographic variables for the professional soccer players (mean ± SD).

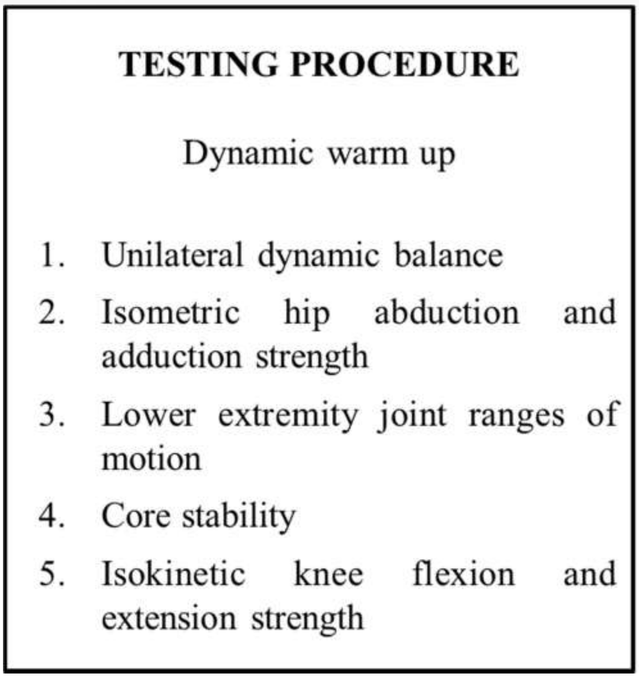
	Males	**Females**
Age (years)	25.5 ± 5.0	20.1 ± 4.2
Stature (cm)	180.1 ± 6.5	161.4 ± 5.2
Body mass (kg)	75.0 ± 6.5	57.2 ± 9.7
Years playing soccer (years)	16.1 ± 4.0	8.4 ± 3.1
Weekly practice frequency (days)	6.1 ± 1.2	3.3 ± 1.4
Hours of soccer practice per week	9.8 ± 2.1	5.1 ± 1.7
Hours of soccer practice per training session	1.6 ± 0.2	1.3 ± 0.3

SD: standard deviation.

7.3.3. Study design and procedure

The relationships of several parameters of neuromuscular performance with UDB measured through Y-BT were determined using a cross sectional and observational study design. Testing was performed during the preseason for both male and female teams, which was at the beginning of August and September, respectively.

Prior to the neuromuscular testing, all participants performed a standardized dynamic warm-up. Three to five minutes after the dynamic warm-up was carried out, participants started five different testing manoeuvres: 1) unilateral dynamic balance; 2) isometric hip abduction and adduction strength; 3) lower extremity joint ranges of motion; 4) core stability; and 5) isokinetic knee flexion and extension strength. The order of the tests was consistent for all participants with the intention of minimizing any possible negative influence among variables (Figure 7.1).



TESTING PROCEDURE

Dynamic warm up

1. Unilateral dynamic balance
2. Isometric hip abduction and adduction strength
3. Lower extremity joint ranges of motion
4. Core stability
5. Isokinetic knee flexion and extension strength

Figure 7.1. Testing procedure.

Each of the 6 testers who took part in this study conducted the same tests through all the testing sessions and they were blinded to the purposes of this study. All testers had more than 4 years of experience in the neuromuscular assessments.

7.3.3.1. Unilateral dynamic balance

UDB was measured using the Y-BT (Y-Balance Test™, Move2Perform, Evanville, IN), and followed the guidelines proposed by Shaffer et al.[258]

Players were allowed a maximum of five trials to obtain three successful trials for each reach direction (anterior [Figure 7.2a], posteromedial [Figure 7.2b] and posterolateral [Figure 7.2c]). Trials were discarded if the player failed to maintain unilateral stance on the platform, failed to maintain reach foot contact with the reach indicator on the target area while the reach indicator is in motion, used the reach indicator for stance support, or failed to return the reach foot to the starting position under control.[258] Specifically, the testing order was completed as dominant anterior, non-dominant anterior, dominant posteromedial, non-dominant posteromedial, dominant posterolateral, and non-dominant posterolateral. The dominant leg was defined as the participant´s preferred kicking leg (self-reported). The average of the three reaches was normalised by dividing by leg length to standardize the maximum reach distance ([excursion distance/leg length] x100 = % maximum reach distance).[301] Leg length was defined as the length measured in centimetres from the anterior superior iliac spine to the most distal portion of the medial tibial malleolus. To obtain a global measure of the balance performance, data from each direction was averaged to determine a composite score.[311]

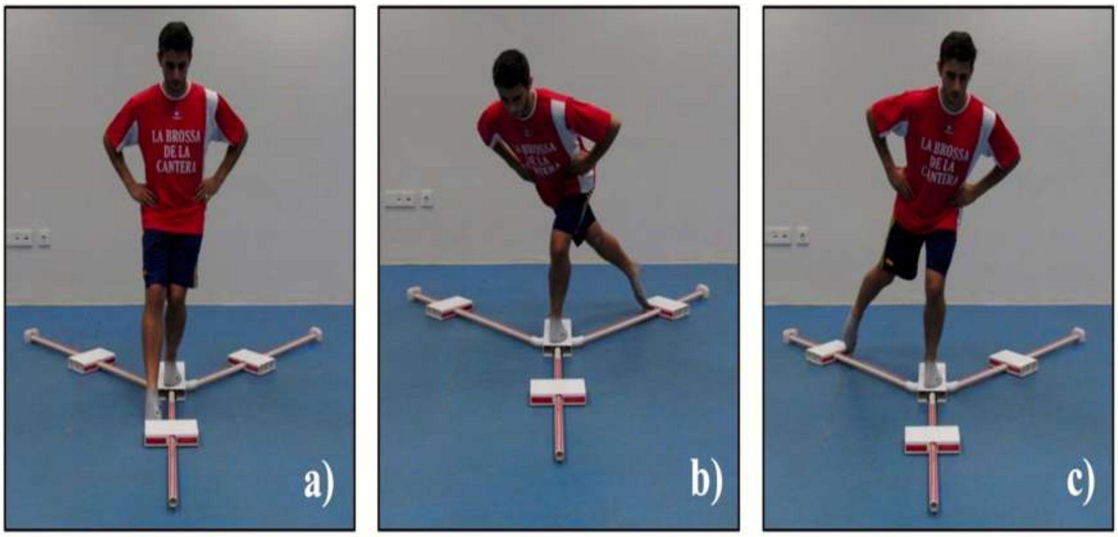


Figure 7.2. Y-Balance Test™ directions; a) anterior reach direction; b) posteromedial reach direction; c) posterolateral reach direction.

7.3.3.2. Isometric hip abduction and adduction strength

Isometric hip abduction and adduction peak torque (PT_{HABD} and PT_{HADD}) of the dominant and non-dominant leg were assessed with a portable handheld dynamometer (Nicholas Manual Muscle Tester, Lafayette Indiana Instruments) in a supine lying position on a plinth with the participants' legs extended (Figure 7.3), following the methods described by Thorborg et al.[259] Briefly, participants performed five trials of 5-seconds isometric maximal voluntary contraction for each hip movement to reduce learning effects. Peak torque values were normalised by body mass. The mean of the three most closely related trials were used for subsequent statistical analyses.

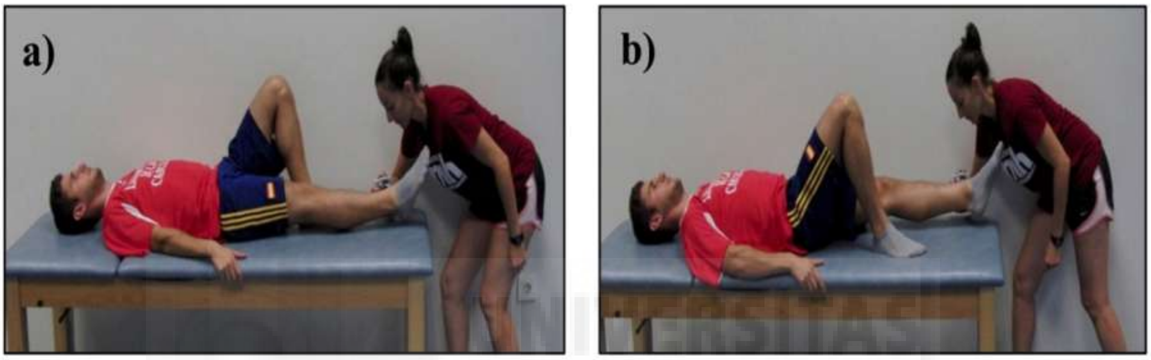


Figure 7.3. Isometric hip adduction (a) and abduction (b) strength assessment.

7.3.3.3. Lower extremity joint ranges of motion

Passive hip flexion with knee flexed (PHF_{KF}) (Figure 7.4a) and extended (PHF_{KE}) (Figure 7.4b), extension (PHE) (Figure 7.4c), abduction (PHA) (Figure 7.4d), external (PHER) (Figure 7.4e) and internal (PHIR) (Figure 7.4f) rotation; knee flexion (PKF) (Figure 7.4g); and ankle dorsiflexion with knee flexed (ADF_{KF}) (Figure 7.4h) and extended (ADF_{KE}) (Figure 7.4i) ROMs of the dominant and non-dominant leg were assessed following the methods previously described.[224] Two maximal trials for each leg were carried out, and the mean score for each test was used in the subsequent analyses. When a variation > 5% was found in the ROM values between the two trials of any test, an extra trial was performed.

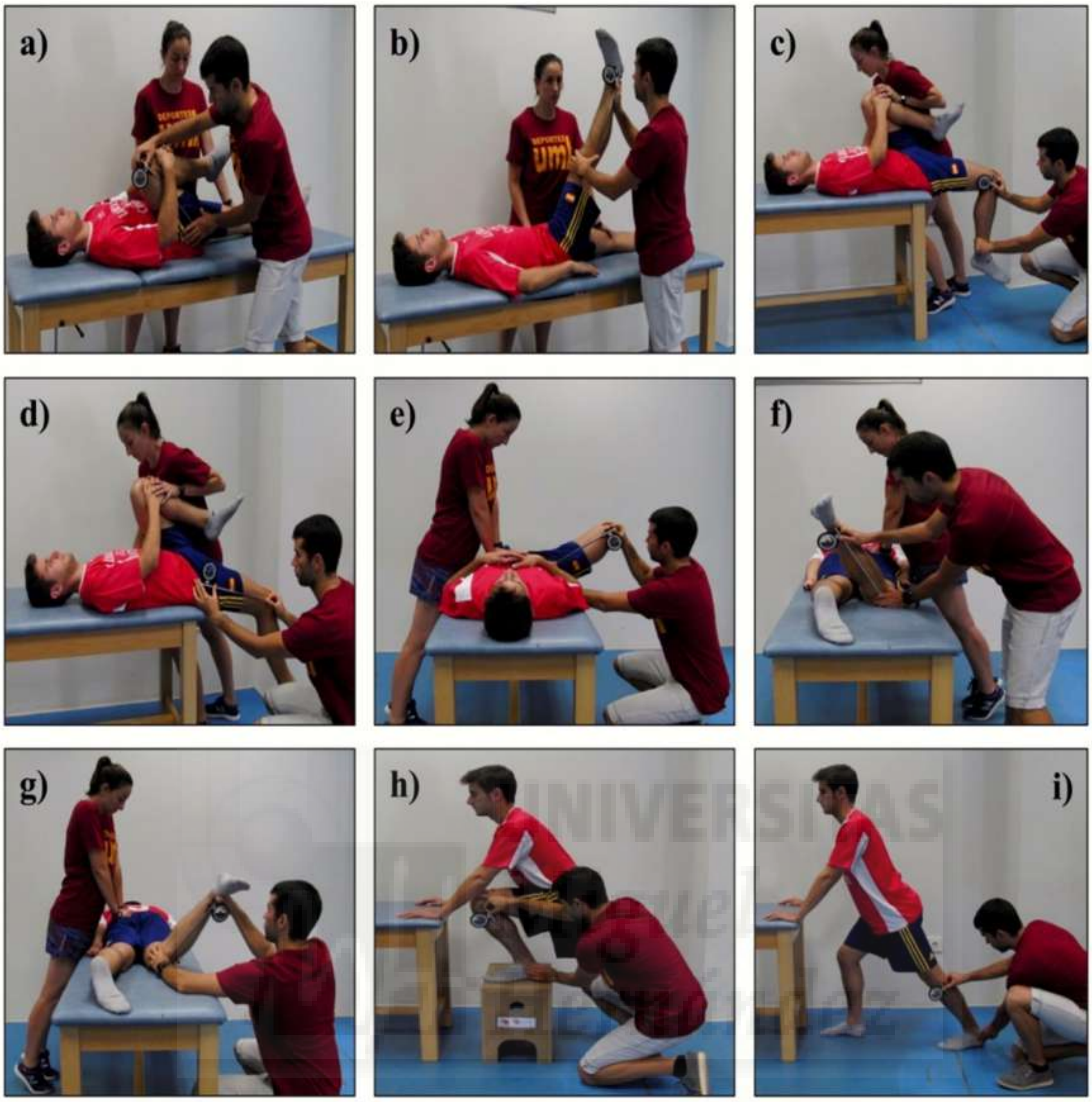


Figure 7.4. Lower extremity joint ranges of motion assessment: a) passive hip flexion with knee flexed test [PHF_{KF}]; b) passive hip flexion with knee extended test [PHF_{KE}]; c) passive knee flexion [PKF]; d) passive hip extension [PHE]; e) passive hip abduction [PHA]; f) passive hip external rotation test [PHER]; g) passive hip internal rotation test [PHIR]; h) ankle dorsiflexion with knee flexed test [ADF_{KF}]; and i) ankle dorsiflexion with knee extended test [ADF_{KE}].

7.3.3.4. Core stability

An unstable sitting protocol was used to assess participants' CS (defined as the ability to control trunk posture and motion while sitting, following the methods previously described by Barbado et al.[260] Participants performed two static and three dynamic trials while sitting on an unstable seat (Figure 7.5) which was placed on a force plate (Kistler, Switzerland, Model 9286AA). One of the static trials was performed without visual feedback (CS_{NF}), in which participants were asked to sit as still as possible in their preferred seated position; and the other static trial was performed with visual feedback (CS_{WF}), in which participants were requested to adjust their CoP position to a target point located in the centre of a screen placed in front of the them. The three dynamic trials were executed with visual feedback. In these trials, participants were asked to track the target point which moved along three possible trajectories: medial-lateral (CS_{ML}), anterior-posterior (CS_{AP}), circular (CS_{CD}). Feedback of the CoP and target point displacement was provided to the participants in real time (Figure 7.5).

The duration of each trial was 70 seconds and the rest period between trials was 1 minute. The full protocol was performed twice and 2 minutes of practice was given to participants before recording.

To quantify the trunk control during the sitting trials, we used the mean radial error (MRE). MRE was calculated as the average of vector distance magnitude (mm) of the CoP from the target point or from the participant's own mean CoP position[261] for trials with and without visual feedback, respectively. The best of two trials performed for each condition (lower MRE) was used for subsequent statistical analyses.

Figure 7.5. Participant performing unstable sitting protocol. Projection providing visual feedback of participants' centre of pressure and a target point moving across a circular path.

7.3.3.5. Isokinetic knee flexion and extension strength

A Biodex System-4 isokinetic dynamometer (Biodex Corp., Shirley, NY, USA) and its respective manufacture software were used to determine isokinetic concentric and eccentric torques during knee extension and flexion actions in both legs.

Participants were secured in a supine position with the hip passively flexed at 10°–20° (Figure 7.6). The axis of rotation of the dynamometer lever arm was aligned with the lateral epicondyle of the knee. The force pad was placed approximately 3 cm superior to the medial malleolus, with the foot in a relaxed position. Adjustable strapping across the pelvis, thigh proximal to the knee and foot localized the action of the musculature involved. The range of movement was set from 90° knee flexion (starting position) to 0° (0° was determined as maximal voluntary knee extension for each participant).

Before isokinetic testing, the participants performed a specific isokinetic warm-up consisting of three sub-maximal (self-perceived 50% effort) and two maximal concentric and eccentric knee extension and flexion actions at 120°/s.

The isokinetic examination was separated into two parts. The first part of the examination was the assessment of the knee extensor and knee flexor muscles during concentric/concentric (CON/CON) cycles with extension undertaken first. After a 5 min rest period the eccentric/eccentric (ECC/ECC) testing cycle was performed. In both testing methods, three sets of two cycles of knee flexion and extensions were performed at three present constant angular velocities in the following order: 60, 180, 240 and 300°/s for CON/CON cycles; and 30, 60 and 180º/s for ECC/ECC cycles (slow to fast). The two testing parts (CON/CON and ECC/ECC) were separated by a 5 min rest interval and a rest of 30 s was allowed between action cycles. For both concentric and eccentric actions, participants were encouraged to push–pull/resist as hard and as fast as possible and to complete the full range of motion. Participants were instructed to abort the test if they felt any discomfort or pain. During the test, all participants were given visual feedback from the system monitor. They were also verbally encouraged by the investigator to give their maximal effort, and the instructions were standardized by using keywords such as ‘resist’, ‘push’ and ‘hard and fast as possible’.

The isokinetic gravity-corrected and normalised to body weight peak torque (PT) variable was extracted for each movement (flexion and extension), muscle action (concentric, eccentric) and velocity (60, 180 and 240º/s for concentric actions and 30, 60 and 180º/s for eccentric actions). In each of the three trials at each velocity, the PT was reported as the single highest torque output. The average PT score of the 3 sets at each velocity was used for subsequent statistical analysis. When a variation > 5% was found in the PT values between the three trials, the mean of the two most closely related torque values were used for the statistical analyses.

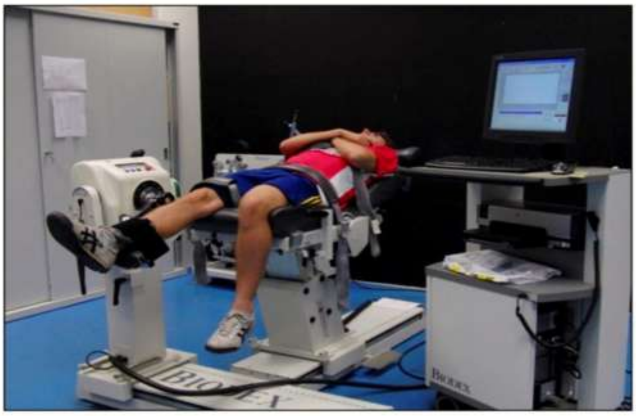


Figure 7.6. Isokinetic knee flexion and extension strength assessment.

7.3.4. Statistical analysis

The distributions of raw data sets were checked using the Kolomogorov-Smirnov test and demonstrated that all data had a normal distribution ($p > 0.05$). Descriptive statistics including means and standard deviations were calculated for each variable. Dependent t tests were used to test for differences between the scores of the dominant and non-dominant leg for the measures obtained from the unilateral testing manoeuvres (unilateral dynamic balance; isometric hip abduction and adduction strength; lower extremity joint ranges of motion; and isokinetic knee flexion and extension strength). Independent t tests were applied to examine sex differences in the neuromuscular parameters.

Pearson’s correlation moments (r) were used to assess the relationship between trunk stability, range of motion, hip and knee strength and Y-Balance performance. Subsequently, in order to identify a group of factors that were independently associated with balance performance, all potential factors that showed significant associations with the composite normalised reach score of Y-BT and met the assumptions of normality, linearity, homoscedasticity, and presence of multicollinearity were entered into a stepwise multivariate linear least square regression with backward elimination ($p \leq 0.10$). Normality, linearity, and homoscedasticity assumptions were confirmed via observation of both the normality probability plots of the regression standardized residual plots and the standardized residual versus the regression standardized predicted value scatterplots. Multicollinearity was defined as a Pearson product moment correlation coefficient between 2 variables of equal to or greater than 0.7, therefore only those parameters which showed correlations lower than 0.7 was used for further analysis. The strength of the predictive ability of identified factors was determined with unstandardized regression coefficients (β), while the predictive power of each final model was given by calculation of the percentage of explained variance (R^2). Both, correlational and multiple regression analysis were performed for males and females independently. Potential confounding variables (age, mass, stature, playing experience and competitive level) were included in the regression model. Significance level was set at a level of $p < 0.05$. All data were analysed using the Statistical Package for Social Sciences (version 22 for Windows, SPSS Inc, Chicago, IL, USA).

7.4. Results

Descriptive statistics (mean ± standard deviation) for each variable are displayed in table 7.2. Dependent t-test analysis reported statistically significant differences ($p < 0.05$) between dominant and non-dominant legs for several isokinetic knee flexion and extension PT (concentric and eccentric) values and some joint ROMs in males; whereas females only showed between legs differences for two isokinetic eccentric PT values of the knee extensors (PTs measured at 30 and 60º/s) and three ROMs (PKF, ADF_{KF}, ADF_{KE}). Independent t test analysis showed sex-related differences in most isokinetic and isometric strength measures, whereby males reported higher scores than females. In addition, males reported statistically significant lower PHF_{KF}, PHER, PHIR, PHE and ADF_{KF} ROMs than females.

All neuromuscular variables that reported statistically significant correlations with the composite score of the Y-BT (in both sexes) and containing no intercorrelations above 0.7 were used for multivariate regression analysis (Tables 7.3 and 7.4 for males and females respectively). The composite scores of dominant and non-dominant legs were considered as the class or criterion variable. Mass was introduced as a potential confounder parameter for females. As it can be showed in table 7.5, PHF_{KF} and ADF_{KF} ROMs were the main factors that demonstrated a significant association with dominant and non-dominant composite scores for males. The model derived for the non-dominant leg composite score showed a greater degree of explained variance (33.5%) than for dominant leg (23.1%).

Table 7.2. Neuromuscular performance values of male and female professional soccer players (mean ± SD).

Variable	Males		Females	
	Dominant leg	Non-dominant leg	Dominant leg	Non-dominant leg
Unilateral dynamic balance				
Composite	88.3 ± 7.8	88.7 ± 7.1	86.8 ± 6.4	87.7 ± 5.6
*Isokinetic strength (N*m/kg)*				
Concentric KF*				
▪ PT$_{60}$	1.18 ± 0.24^{T}	1.10 ± 0.22	0.85 ± 0.17	0.82 ± 0.17
▪ PT$_{180}$	0.95 ± 0.22	0.92 ± 0.20	0.57 ± 0.19	0.61 ± 0.19
▪ PT$_{240}$	0.90 ± 0.19^{T}	0.86 ± 0.18	0.58 ± 0.21	0.59 ± 0.20
▪ PT$_{300}$	0.86 ± 0.19^{T}	0.78 ± 0.18	0.58 ± 0.18	0.56 ± 0.19
Concentric KE*				
▪ PT$_{60}$	2.49 ± 0.48^{T}	2.39 ± 0.44	1.73 ± 0.39	1.75 ± 0.41
▪ PT$_{180}$	1.69 ± 0.32	1.69 ± 0.33	1.07 ± 0.34	1.09 ± 0.39
▪ PT$_{240}$	1.51 ± 0.30	1.49 ± 0.33	1.00 ± 0.33	1.00 ± 0.34
▪ PT$_{300}$	1.39 ± 0.27	1.34 ± 0.32	0.97 ± 0.30	0.92 ± 0.30
Eccentric KF				
▪ PT$_{30}$	2.70 ± 0.83^{T}	2.56 ± 0.80	2.64 ± 0.64	2.53 ± 0.65
▪ PT$_{60}$	2.78 ± 0.74^{T}	2.62 ± 0.79	2.54 ± 0.65	2.57 ± 0.61
▪ PT$_{180}$	2.52 ± 0.71	2.43 ± 0.77	2.36 ± 0.56	2.43 ± 0.56
Eccentric KE				
▪ PT$_{30}$	1.20 ± 0.30^{T}	1.08 ± 0.25*	1.10 ± 0.26^{T}	0.98 ± 0.20
▪ PT$_{60}$	1.21 ± 0.29*^{T}	1.09 ± 0.26*	1.06 ± 0.25*^{T}	0.98 ± 0.18
▪ PT$_{180}$	1.19 ± 0.29^{T}	1.09 ± 0.27*	1.07 ± 0.27	0.99 ± 0.27
Isometric hip strength (N/kg)				
PT$_{HABD}$*	2.71 ± 0.40	2.77 ± 0.39	2.38 ± 0.41	2.44 ± 0.39
PT$_{HADD}$*	2.74 ± 0.53	2.69 ± 0.47	2.30 ± 0.44	2.28 ± 0.36
Core stability (mm)				
CS$_{NF}$	6.11 ± 2.16*		4.31 ± 1.71	
CS$_{WF}$	5.34 ± 1.44		5.48 ± 2.43	
CS$_{ML}$	8.27 ± 2.01*		7.19 ± 2.49	
CS$_{AP}$	8.30 ± 1.70*		7.24 ± 2.11	
CS$_{CD}$	10.79 ± 2.96*		9.22 ± 3.75	
Lower extremity joint ROM (º)				
PHF$_{KE}$	80.3 ± 10.7	80.9 ± 10.9	81.6 ± 12.6	81.7 ± 12.5
PHF$_{KF}$	146.6 ± 8.4*^{T}	147.9 ± 7.6*	153.8 ± 8.5	154.5 ± 7.1
PHA	63.9 ± 8.9^{T}	61.5 ± 8.7	63.6 ± 6.9	61.9 ± 7.8
PHER	50.0 ± 9.3*	50.4 ± 9.6*	61.6 ± 6.9	61.2 ± 8.2
PHIR	47.4 ± 8.3*^{T}	45.7 ± 7.7*	56.1 ± 9.1	55.1 ± 8.2
PHE	9.6 ± 8.7*	10.4 ± 8.4*	15.6 ± 5.7	15.6 ± 5.4
PKF	127.6 ± 13.3^{T}	125.5 ± 13.6	130.0 ± 13.8^{T}	129.4 ± 13.6
ADF$_{KF}$	37.2 ± 6.6*^{T}	38.2 ± 5.9	39.9 ± 4.9^{T}	38.1 ± 5.5
ADF$_{KE}$	36.2 ± 5.6	36.5 ± 5.6*	36.0 ± 4.9^{T}	32.7 ± 4.4

*: significant differences between sexes ($p<0.05$); ^{T}: significant differences between legs ($p<0.05$). kg: kilograms; m: metres; mm: millimetres; N: newtons; KE: knee extension; KF: knee flexion; HABD: hip abduction; HADD: hip adduction; PT: peak torque; CS$_{NF}$: unstable sitting without feedback; CS$_{WF}$: unstable sitting with feedback; CS$_{ML}$: unstable sitting while performing medial-lateral displacements with feedback; CS$_{AP}$: unstable sitting while performing anterior-posterior displacements with feedback; CS$_{CD}$: unstable sitting while performing circular displacements with feedback; ROM: range of motion; PHF$_{KE}$: passive hip flexion with knee extended ROM; PHF$_{KF}$: passive hip flexion with knee flexed ROM; PHA: passive hip abduction ROM; PHER: passive hip external rotation ROM; PHIR: passive hip internal rotation ROM; PHE: passive hip extension ROM; PKF: passive knee flexion ROM; ADF$_{KF}$: ankle dorsiflexion with knee flexed ROM; ADF$_{KE}$: ankle dorsiflexion with knee extended ROM.

Table 7.3. Correlations of isokinetic strength of the knee (flexion and extension), isometric strength of the hip (abduction and adduction), core stability and lower extremity joint range of motion measures with composite scores of the dominant and non-dominant leg for stance during Y-Balance test in male professional soccer players (n = 86).

Measure	Unilateral Dynamic Balance	
	Dominant leg	**Non-dominant leg**
*Isokinetic strength (N*m/kg)*		
Concentric KF		
▪ PT_{60}	0.163	0.016
▪ PT_{180}	0.255*	0.013
▪ PT_{240}	0.210	0.151
▪ PT_{300}	0.216	0.194
Concentric KE		
▪ PT_{60}	0.154	0.025
▪ PT_{180}	0.188	0.063
▪ PT_{240}	0.147	0.084
▪ PT_{300}	0.173	0.162
Eccentric KF		
▪ PT_{30}	0.115	0.125
▪ PT_{60}	0.115	0.152
▪ PT_{180}	0.132	0.253*
Eccentric KE		
▪ PT_{30}	0.103	0.067
▪ PT_{60}	0.258*	0.150
▪ PT_{180}	0.352**	0.210
Isometric hip strength (N/kg)		
PT_{HABD}	0.269*	0.135
PT_{HADD}	0.000	0.114
Core stability (mm)		
CS_{NF}	0.054	0.008
CS_{WF}	-0.003	-0.014
CS_{ML}	0.064	-0.024
CS_{AP}	-0.021	-0.052
CS_{CD}	0.026	-0.028
Lower extremity joint ROM (º)		
PHF_{KE}	0.012	0.053
PHF_{KF}	0.382*	0.445*
PHA	-0.035	0.172
PHER	0.021	0.028
PHIR	0.030	0.108
PHE	0.063	0.187
PKF	0.138	0.302**
ADF_{KF}	0.344**	0.429**
ADF_{KE}	0.229*	0.385**

kg: kilograms; m: metres; mm: millimetres; N: newtons; KE: knee extension; KF: knee flexion; HABD: hip abduction; HADD: hip adduction; PT: peak torque; CS_{NF}: unstable sitting without feedback; CS_{WF}: unstable sitting with feedback; CS_{ML}: unstable sitting while performing medial-lateral displacements with feedback; CS_{AP}: unstable sitting while performing anterior-posterior displacements with feedback; CS_{CD}: unstable sitting while performing circular displacements with feedback; ROM: range of motion; PHF_{KE}: passive hip flexion with knee extended ROM; PHF_{KF}: passive hip flexion with knee flexed ROM; PHA: passive hip abduction ROM; PHER: passive hip external rotation ROM; PHIR: passive hip internal rotation ROM; PHE: passive hip extension ROM; PKF: passive knee flexion ROM; ADF_{KF}: ankle dorsiflexion with knee flexed ROM; ADF_{KE}: ankle dorsiflexion with knee extended ROM.*: $p < 0.05$; **: $p < 0.01$.

Table 7.4. Correlations of isokinetic strength of the knee (flexion and extension), isometric strength of the hip (abduction and adduction), core stability and lower extremity joint ranges of motion measures with composite scores of the dominant and non-dominant leg for stance during Y-Balance test in female professional soccer players (n = 44).

Measure	Unilateral Dynamic Balance	
	Dominant leg	**Non-dominant leg**
*Isokinetic strength (N*m/kg)*		
Concentric KF		
▪ PT_{60}	0.086	0.016
▪ PT_{180}	0.076	0.013
▪ PT_{240}	0.153	0.151
▪ PT_{300}	-0.214	0.194
Concentric KE		
▪ PT_{60}	0.379*	0.025
▪ PT_{180}	0.359	0.063
▪ PT_{240}	0.320	0.084
▪ PT_{300}	0.289	0.162
Eccentric KF		
▪ PT_{30}	0.187	0.163
▪ PT_{60}	0.180	0.220
▪ PT_{180}	0.204	0.415*
Eccentric KE		
▪ PT_{30}	-0.021	0.103
▪ PT_{60}	-0.042	0.248
▪ PT_{180}	0.055	0.275
Isometric hip strength (N/kg)		
PT_{HABD}	0.415**	0.529**
PT_{HADD}	0.322*	0.411**
Core stability (mm)		
CSNF	-0.246	-0.386**
CSWF	-0.257	-0.317*
CSML	-0.446**	-0.551**
CSAP	-0.292	-0.423**
CSCD	-0.289	-0.478**
Lower extremity joints ROM (º)		
PHF_{KE}	0.295	0.184
PHF_{KF}	-0.179	-0.212
PHA	0.469**	0.071
PHER	-0.143	-0.129
PHIR	0.297	0.286
PHE	0.399*	0.231
PKF	0.140	0.348*
ADF_{KF}	0.340	0.270
ADF_{KE}	0.284	0.090

kg: kilograms; m: metres; mm: millimetres; N: newtons; KE: knee extension; KF: knee flexion; HABD: hip abduction; HADD: hip adduction; PT: peak torque; CS_{NF}: unstable sitting without feedback; CS_{WF}: unstable sitting with feedback; CS_{ML}: unstable sitting while performing medial-lateral displacements with feedback; CS_{AP}: unstable sitting while performing anterior-posterior displacements with feedback; CS_{CD}: unstable sitting while performing circular displacements with feedback; ROM: range of motion; PHF_{KE}: passive hip flexion with knee extended ROM; PHF_{KF}: passive hip flexion with knee flexed ROM; PHA: passive hip abduction ROM; PHER: passive hip external rotation ROM; PHIR: passive hip internal rotation ROM; PHE: passive hip extension ROM; PKF: passive knee flexion ROM; ADF_{KF}: ankle dorsiflexion with knee flexed ROM; ADF_{KE}: ankle dorsiflexion with knee extended ROM.*: $p < 0.05$; **: $p < 0.01$.

For females (Table 7.5), CS_{ML} and PT_{HABD} demonstrated a significant association for both dominant and non-dominant legs. PHA and AD_{KF} ROMs also showed a significant association for the dominant and non-dominant leg respectively. Similar to males, the model derived for the non-dominant leg composite score explained more of the variance (46.9%) than the dominant leg (38.2%). Overall both derived models for the dominant and non-dominant leg displayed greater predictive power for the females compared with the males.

Table 7.5. Backward multivariate linear regression analysis. Significant predictor variables ($p \leq 0.10$) for the composite normalised reach scores obtained from Y-Balance test.

		Explained variance (R^2)			**Regression equation**
	Model	**1st Variable**	**2nd Variable**	**3rd Variable**	
MALES					
Dominant leg	*23.1%*	PHF_{KF} 14.6%	ADF_{KF} 8.5%		$Y = 33.472 + 0.289*PHF_{KF} + 0.321*ADF_{KF}$
Non-Dominant leg	*33.5%*	ADF_{KF} 20.7%	PHF_{KF} 12.8%		$Y = 24.458 + 0.430* ADF_{KF} + 0.319* PHF_{KF}$
FEMALES					
Dominant leg	38.2%	PHA 22.0%	PT_{HABD} 11.4%	CS_{ML} 4.8%	$Y = 61.632 + 0.311*PHA + 4.258* PT_{HABD} - 0.632* CS_{ML}$
Non-Dominant leg	46.9%	CS_{ML} 31.1%	PT_{HABD} 10.4%	ADF_{KF} 5.4%	$Y = 71.084 - 0.807*CS_{ML} + 5.467* PT_{HABD} + 0.244*ADF_{KF}$

CS_{ML} (mm): core stability during unstable sitting while performing medial-lateral displacements with feedback; PHA(º): passive hip abduction range of motion; PHF_{KF}(º): passive hip flexion with knee flexed range of motion; ADF_{K}(º): ankle dorsiflexion with knee flexed range of motion; PT_{HABD} (N/kg): peak of force during hip abduction exertions; Y: composite normalised reach scores obtained from Y-Balance test.

7.5. Discussion

For males, the results of the current study showed that only PHF_{KF} and ADF_{KF} ROM were significant predictors in determining a meaningful proportion of the R^2 for the Y-BT (composite score) for both dominant (PHF_{KF} = 14.6%; ADF_{KF} = 8.5%) and non-dominant (PHF_{KF} = 19.8%; ADF_{KF} = 12.5%) legs. The combination of these two predictors accounted for 23.1% and 33.5% of the variance in the composite score of the dominant and non-dominant legs respectively.

These results are in agreement with the findings reported by previous studies,[146,147] although not all,[150] who found that ADF_{KF} ROM accounted for an estimated ≈ 20% of the variance in Y-BT in physically active adults. This finding may support the hypothesis that altered ADF_{KF} ROM might influence UDB via mechanical (due to ligamentous insufficiency) and/or functional instability (altered neuromuscular control).[312] This hypothesis has been recently reinforced by data demonstrating that individuals with chronic ankle instability have lower reach distances during the Y-BT when compared to their healthy control counterparts.[302,313]

Regarding PHF_{KF}, Robinson, & Gribble[147] found that hip flexion ROM of the stance leg accounted for approximately 80% of the variance in the Y-BT, which is significantly higher than the R^2 values found in our study (14.6% - 19.8%). A possible explanation for this discrepancy in the magnitude of the explained variance reported by the hip flexion ROM measure with respect to the score achieved in the Y-BT might be based on the fact that Robinson, & Gribble[147] assessed the hip flexion through an electromagnetic tracking system (and its respective kinematic software) while participants performing the Y-BT. Although a priori this methodology could be considered as being more ecologically valid, it does not take into account the possible compensatory movements that may appear in the hip (e.g. posterior pelvic tilt, contralateral inclination, abduction) allowing for greater flexion and thus enabling higher Y-BT scores. These compensatory movements might bias the real contribution of the hip flexion ROM on the Y-BT score. Therefore, in order to carry out an accurate assessment of hip flexion, whilst minimizing the contribution of any compensatory movement, the methodology suggested by Cejudo et al.[224] was followed in the current study. From a theoretical point of view, soccer players with limited hip flexion ROM on the stance leg might show a sub-optimal UDB while performing explosive actions (e.g. kicking and changes of direction) due to less anterior displacement of their center of mass, which may increase the likelihood of loosing stability.

Contrary to the results mentioned above, the measures related to CS, isokinetic strength (concentric and eccentric) of the knee flexors and extensors, isometric strength of the hip adductors and abductors, and ROM of the hip (extension, internal and external rotations, flexion with knee extended) and knee (flexion) joints showed no significant contributions in the Y-BT scores for the dominant and non-dominant legs in males soccer players. The results found in this study for the isokinetic strength of the knee are in concordance with the findings reported by Booysen, Gradidge, & Watson[141] and Lockie, Schultz, Callaghan, Jeffriess, & Berry,[142] who did not show any relationship between the isokinetic strength of the knee flexors and extensors and the Y-BT score in professional soccer players and team-sport athletes respectively. However, conflicting data are available where knee strength has been a significant predictor of Y-BT score, and consequently in UDB, in male amateur soccer players[141,143] and recreational athletes.[314] Perhaps, knee strength might be a limiting factor of UDB in athletes who show low to moderate strength scores while in athletes with high strength scores, this parameter may cease to be relevant. The other variables measured cannot be compared with previous studies because this is the first that has determined these neuromuscular parameters in relation to UDB. Thus, the findings of the current study highlight the meaningful role of the hip and ankle ROM in the sagittal plane and the strength of the hip abductors on the UDB in professional male soccer players. Consequently, training interventions aimed at improving or maintaining UDB in professional male soccer players should include, among others, stretching exercises for the posterior chain of the lower extremity.

For females, the findings of the current study demonstrate that PHA (22%), PT_{HABD} (11.4%) and CS_{ML} (4.8%) are all significant predictors in Y-BT for the dominant leg, whereas the CS_{ML} (31.1%), PT_{HABD} (10.4%) and ADF_{KF} (5.4%) are significant predictors for the non-dominant leg. The total explained variance values of these predictors in the composite scores of the Y-BT were 38.2% and 46.9% for the dominant and non-dominant leg respectively.

Similar results than those found in the current study for the PHA and ADF_{KF} measures have been reported in previous studies[146,147] indicating that individuals with higher scores in such variables achieved better performance in the Y-BT and therefore demonstrate superior UDB. Previously correlations between PT_{HABD} and the distances achieved in the Y-BT have been demonstrated by Ambegaonkar, Mettinger, Caswell, Burtt, & Cortes[144] in university lacrosse and soccer players ($R^2 = 11.5\%$). The positive and significant relationship found in the current study between hip abductor strength and UDB may be attributed to the role that the abductor muscles have shown during single leg movements,[315] as a hip stabiliser in the frontal plane (reducing knee adduction moment), which might reduce the likelihood of sustaining lower extremity injuries (mainly ACL tears).

Although CS has been proposed as a crucial factor for UDB,[315] this is the first study (to the authors´ knowledge) confirming this link in professional female soccer players. Deficits in CS, especially in the frontal plane, could lead to uncontrolled upper body displacements during single leg movements, moving away the center of mass of the body from the foot of support, which might compromise dynamic stability of the lower limb.[316] This comprised UDB may predisposed to excessive motion in the hip or trunk, potentially permitting their entire lower extremity to move into positions frequently associated with noncontact injuries such as femoral adduction and internal rotation,[315] which in turn, may increase knee injury risk during common soccer actions such as landing and cutting.[317]

Contrary to the results mentioned above, and similar to males, the rest of the measures related to CS (with the exception of CS_{ML}), isokinetic strength (concentric and eccentric) of the knee flexors and extensors, isometric strength of the hip adductors, and ROM of the hip (extension, internal and external rotations, flexion with knee extended) and knee (flexion) joints showed no significant contributions in the Y-BT scores for the dominant and non-dominant legs.

Therefore, and in order to improve or maintain UDB, training intervention programs in females soccer players should be focused on exercises designate to improve: a) ankle dorsiflexion and hip abduction ROMs; b) strength and mobility of the hip abductors; and c) CS (especially in the frontal plane).

For both sexes, the regression equations generated only explained modest percentages of the performance achieved in the composite score of the Y-BT (23.1% [dominant leg] and 33.5% [non-dominant leg] for males and 38.2% [dominant leg] and 46.9% [non-dominant leg] for females). Comparisons with other regression models were not possible as this is the first study that have analysed the concurrent influence of a range of training modifiable neuromuscular measures in the Y-BT performance. Future studies should consider the inclusion of other factors such as core endurance, muscle stiffness and closed chain lower extremity strength measures in the regression analysis in order to determine whether they would increase the modest percentages of explained variance reported in this study.

On the other hand, the current study is the first that have analysed the possible sex-related differences in UDB reached distances in soccer players as well as if males and females used similar or different neuromuscular strategies to achieve them. Thus, the findings of this study indicate that male and female professional soccer players have similar UDB scores measured through the Y-BT. Analogous results in other cohorts have been found by previous researchers that have demonstrated no difference between sexes in performance on the YBT in Collegiate Athletic Association Division I athletes,[148] basketball players[149] and recreational athletes.[150] Furthermore, in the present investigation, there was no limb differences in Y-BT reach performance for either males or females, which is consistent with the finding reported by one previous study.[143] Likewise, the Y-BT reach scores were highly comparable with the results reported by Booysen et al.,[141] in male soccer players, although higher than values achieved by active adults.[143,147] These results indicate that the Y-BT may be sensitive to level of training and/or sport-related adaptations.[143,307,308] Finally, the current findings also suggest that there were sex-related differences in the isokinetic knee (flexion and extension) and isometric hip (abduction and adduction) strength values and joint ROMs. Previous investigations have not reported knee and hip strength values for both limbs in soccer players, thus it is difficult to draw comparisons with the current findings. However, the sex-related differences found were within expected ranges for healthy active adults.[318]

However, and despite the fact that male and female professional soccer players performed similarly on the Y-BT, the overall balance score was achieved in different ways. Thus, for males, those variables related to movement patterns in the sagittal plane (PHF_{KF} and ADF_{KF} ROM measures) are important in the composite score obtained in the Y-BT, with an overall explained variance of 23.1% and 32.3% for the dominant and non-dominant legs respectively. However, for females, variables related to the performance of movement patterns in the frontal plane such as core stability (CS_{ML}), hip abduction strength (PT_{HABD}) and ROM (PHA) were considered predictors of Y-BT reached distances, accounting together for 38.2% and 41.5% of the variance in the composite score of the Y-BT. These findings appear to suggest a sex-specific related injury profile when professional soccer players report poor UDB values. For example, males might be more prone to suffer ankle injuries due to mechanical and/or neuromuscular instability in the joint; while women would be more likely to adopt inappropriate movement strategies in the frontal plane (e.g. dynamic knee valgus) during landing and cutting tasks, leading to a higher risk of knee injury (mainly ACL tears and patella femoral pain). This assumption, if true, may partially explain the fact that females have a 3-8 times greater risk of ACL injury than similarly trained males.[319] This may be important for female soccer players who are at greater risk of ACL injury than males due to non-modifiable (anatomical, hormonal) and modifiable (neuromuscular) risk factors especially in sports that require cutting and landing motions, such as soccer.[320]

7.5.1. Limitations

The authors acknowledge several study limitations. The current findings are limited to our participants' sport background (professional soccer players) so that extrapolation to other sport cohorts should be made with a certain degree of caution. Another limitation of the study is that there were a disproportionate number of male participants compared to female participants. However, it should be noted that the variability in the male and female scores on the Y-BT and neuromuscular parameters measured were similar. This observation would suggest that any Type II error that may be attributed to lower participant numbers and high levels of variability would be minimal due to the relatively normalised variance between the sexes. Isokinetic (knee) and isometric (hip) strength was tested in an open chain rather than in close chain movement. This may have resulted in the non-significant correlations observed in both males and females, due to the lack of movement specificity between open chain dynamometry and functional performance. However, the researchers felt that this was the most objective and reliable way of assessing hip and knee strength.

7.6. Conclusions

The main findings of the current study indicate that despite the fact that male and female professional soccer players reported similar UDB scores, different measures of neuromuscular performance appear to have influence this fundamental ability. Thus, for males, those variables related to movement patterns in the sagittal plane (PHF_{KF} and ADF_{KF} ROM measures) are important in the overall balance score obtained in the Y-BT. However, for females, variables related to the performance of movement patterns in the frontal plane such as core stability (CS_{ML}), hip abduction strength (PT_{HABD}) and ROM (PHA) were considered predictors of Y-BT reached distances. Therefore, and in order to design balance-training interventions aimed at improving or maintaining UDB in professional soccer players, for males should include, among other, stretching exercises for the posterior chain of the lower extremity; whereas for females should be focused on exercises designate to improve ankle dorsiflexion ROM, strength and mobility of the hip abductors and CS (especially in the frontal plane).

CHAPTER 8

EPILOGUE

CHAPTER 8

EPILOGUE

8.1. General conclusions

The studies included in this doctoral thesis: a) provide a unique understanding of the extent of the injury problem in male professional soccer players; b) present two original injury risk factor-based models to identify players at high or low risk of lower-extremity muscle injuries and hamstring strains by applying a novel multifactorial approach; and c) improve the knowledge base regarding to the relationship between some of the most important soccer risk factors.

The first study explored and analysed the extent of the injury problem in male professional soccer players through a systematic review and a novel meta-analysis of the literature. On the other hand, the second study described lower extremity range of motion profile in professional soccer players and analysed differences between goalkeepers and outfield players. The third and fourth studies developed two models to identify professional soccer and handball players at high or low risk of lower extremity muscle and hamstring injuries by applying several pre-processing, cost-sensitive learning and ensemble techniques from machine learning. The last study examined the relationships between several parameters of neuromuscular performance with unilateral dynamic balance and determined sex-related differences in the influence of neuromuscular parameters on dynamic balance.

The main findings obtained in this doctoral thesis may help clinicians and sport practitioners to have a greater understanding of the etiology of soccer-related injuries, and the interactions between different risk factors and a broader approach to explore the cause of injury. In addition, they might use the models built to identify professional soccer and handball players at high or low risk of lower extremity muscle and hamstring injuries during preseason screenings. Furthermore, the findings of the studies two and five of this thesis may also help professionals to understand the possible soccer-specific adaptations in the lower extremity joint ranges of motion better and to individualize dynamic balance training interventions according to sex-related differences. Overall, this thesis provides new information to facilitate the decision-making process to choose the best strategies for injury prevention.

The following summarizes the major contributions of this thesis:

Study 1:

1. Professional male soccer players are exposed to a substantial risk of sustaining injuries, especially during matches.
2. The lower extremity was the most frequently injured body region, being the thigh the anatomical region in which injuries occurred more habitually.
3. The most common type of injuries was muscle/tendon strains.
4. Most injuries appeared to be of minimal severity and caused by a traumatic mechanism.

Study 2:

5. A high number of the soccer players presented restricted passive hip flexion with knee extended and/or ankle dorsiflexion with knee flexed range of motion values.
6. There were no significant differences in range of motion measures for the hip, knee and ankle between outfield players and goalkeepers.
7. Some bilateral differences were identified for passive hip abduction, passive hip internal and external rotation in professional soccer players.

Study 3:

8. The muscle injury prediction model showed high accuracy for identifying professional soccer and handball players at high or low risk of injury during preseason screenings. Thus, the model reported an area under the curve score of 0.747 with true positive and negative rates of 65.9% and 79.1% respectively. Sport devaluation, history of muscle injury in the previous season and angle peak torque measured through concentric (quadriceps) and eccentric (hamstrings) knee extension movements were identified as the most influential parameters on muscle injury.

Study 4:

9. The hamstring prediction model showed high accuracy for identifying professional soccer and handball players at risk of hamstring injuries during preseason screenings. Thus, the model reported an area under the curve score of 0.867 with false positive and negative rates of 10% and 28% respectively. Sleep quality, reduced sense of accomplishment, history of hamstring strain injury in the previous season, reciprocal hamstring-to-quadriceps ratios and hamstrings-quadriceps torque values obtained close to knee extension were identified as the most influential parameters on hamstring strain injury.

Study 5:

10. Male and female professional soccer players were influenced by different neuromuscular performance variables in their ability to maintain unilateral dynamic balance.
11. Male player stability was determined by movement patterns in the sagittal plane, such as passive hip flexion with knee flexed and ankle dorsiflexion with knee flexed ranges of motion.
12. Female player stability was determined by movement patterns in the frontal plane such as core stability (medial-lateral displacements), hip abduction strength and ankle dorsiflexion with knee flexed range of motion.

8.2. Thesis limitations and future research

Most limitations of this doctoral thesis have been addressed in the discussion section of each of the five studies (chapters 3, 4, 5, 6, 7). In addition, this section presents several limitations that have been the origin of new research projects in the Biomechanics and Health Laboratory of the Sports Research Center of Miguel Hernandez University of Elche. Briefly, the new research purposes are the following:

1. *To collect and analyse injury incidence of different team sports (e.g. basketball, handball, futsal) and levels of play (professionals or amateurs).* In this thesis I have only analysed the epidemiology of injuries in professional male soccer, which has allowed us to know the main characteristics of the injuries and with that to establish prediction models on injuries with the highest incidence rates. However, I cannot claim that these data are the same in other age groups, levels of play or sex, and especially in other team sports. Therefore, it is essential that future studies investigate these aspects in order to develop preventive models according to each population or sport discipline.
2. *To evaluate the applicability of the protocols developed in this doctoral thesis to different populations.* Due to the intrinsic characteristics of soccer, i.e. type and location of injuries with higher incidence, associated risk factors, physical requirements, etc., it is possible that the results obtained in studies three and four might not be generalizable to other populations. Therefore, future lines of work should explore the applicability of these protocols to other high-performance sports with a high incidence of injuries, as well as untrained individuals, different age groups, levels of play and sex. In this case, a doctoral thesis which is being developed in our research group is trying to replicate the current results in other high-performance sports in both sexes.
3. *To include more evidence-based risk factors in the prediction models.* Although in the injury prediction models more than 200 variables related to risk factors described in the literature were analysed, it was not possible to include other risk factors due to methodological and instrumental limitations, which could improve the models obtained. Among these risk factors I could highlight genetic factors[69,70] physical fitness factors such as VO2max[79] or kinematics during jumping and landing.[321] In this sense, some studies of our research group have included these risk factors in their analysis.

4. *To improve complex statistical approaches coming from machine learning and data mining environments used in this thesis.* The mathematical models presented in studies three and four still have some limitations, as having a model with good predictive accuracy is not enough if we are interested in answering why an injury happened and what predictors are associated with it.[8,127] We cannot answer these questions if we use the same approach as we did for resampling techniques as alternating decision trees, since they only allow to dichotomize the player. For example, we might be interested in how much an injury likelihood will increase if the hamstrings strength imbalance between the player's legs increases or if there is a deficit in the range of motion, which could be estimated from statistical models such as Bayesian networks.
5. *To perform field based tests to create prediction models that are more accessible to sport science.* One of the main limitations of this thesis was the technical complexity of the mathematical models used in some studies, as well as the specificity of the laboratory protocols used. Although these attributes give a great methodological value to the doctoral thesis, they make the replication of the main conclusions obtained very difficult. Therefore, our research group is currently working on a line of research focused on the use of field based tests, less expensive, more specific and easier to carry out, with the intention of creating predictive models of injuries closer to the reality of high performance sport. In addition, our group is creating a mobile application which might be easily used in clinical, sport and research settings to predict injuries, as its use does not require knowing complex statistical techniques.
6. *To examine possible interactions between more risk factors.* This thesis has shown the interaction of several neuromuscular parameters (e.g. lower extremity range of motion and core stability) on dynamic unilateral stability. However, very few parameters were included in the regression models, and few studies have investigated interactions among other neuromuscular parameters considered as risk factors in the literature. Further research should investigate the relationships between other neuromuscular parameters and/or include more complex nonlinear statistical models that may show other types of interactions between risk factors.

7. *To check the existence of biomechanical sport-specific adaptations in other populations.* Despite the different physical and physiological demands of soccer, this thesis did not find specific adaptations of the players in the range of motion of the lower extremity with respect to the values published in sedentary population. However, studies in other sports, such as handball, golf or tennis, found specific adaptations in shoulder ranges of motion in comparison with control populations,[249,322,323] while other three studies found differences in hip range of motion in handball and futsal players compared to untrained people.[224,238,324] Future research should confirm these findings in other professional teams and other levels of play before affirming that there are or there are not specific adaptations in the range of motion of the lower extremity in soccer.
8. *To develop prevention protocols once high-risk players are identified.* The injury prevention models establish that once the information about the problem (injury) and the main causes of injury are obtained, the next step is to propose preventive programs that ratify the data obtained and show the effectiveness of preventive strategies. Therefore, another limitation of this thesis, and future line of study, is the need to corroborate in a practical way the main results of these studies, establishing totally individualized preventive programs based on the identification of players with greater injury risk and the risk factors associated with each player.

CHAPTER 9

EPÍLOGO

CHAPTER 9

EPÍLOGO

9.1. Conclusiones generales

Los estudios incluidos en esta tesis doctoral: a) proporcionan una mejor comprensión del alcance del problema de las lesiones en jugadores masculinos de fútbol profesional; b) presentan dos modelos basados en factores de riesgo de las lesiones que identifican a jugadores con alto o bajo riesgo de lesión muscular en la extremidad inferior y/o de desgarro de los isquiosurales aplicando un nuevo enfoque multifactorial, y c) mejoran el conocimiento respecto a la interacción entre algunos de los factores de riesgo más importantes del fútbol.

El primer estudio recopiló y analizó el alcance del problema de las lesiones en jugadores masculinos de fútbol profesional a través de una revisión sistemática y un meta-análisis de la literatura. Por otro lado, el segundo estudio describió el perfil del rango de movimiento de la extremidad inferior en jugadores de fútbol profesional y analizó las diferencias entre porteros y jugadores de campo. El tercer y cuarto estudio desarrollaron dos modelos para identificar jugadores de fútbol y balonmano profesionales con alto o bajo riesgo de lesiones musculares de las extremidades inferiores y de lesiones de los isquiosurales mediante la aplicación de varias técnicas de aprendizaje automático. El último estudio analizó las relaciones entre varios parámetros de rendimiento neuromuscular con la estabilidad dinámica unilateral y determinó diferencias relacionadas con el sexo en la influencia de los parámetros neuromusculares sobre el equilibrio dinámico.

Los principales hallazgos obtenidos en esta tesis doctoral pueden ayudar a los profesionales de la medicina deportiva y a los deportistas a tener una mayor comprensión de la etiología de las lesiones relacionadas con el fútbol, las interacciones entre los diferentes factores de riesgo y un enfoque más amplio que permita explorar la causa de la lesión. Además, podrán usar los modelos diseñados para identificar jugadores profesionales de fútbol y balonmano con alto o bajo riesgo de lesión muscular en las extremidades inferiores o de isquiosurales durante las valoraciones de pre-temporada. Además, los hallazgos de los estudios dos y cinco de esta tesis también pueden ayudar a los profesionales del ámbito futbolístico a comprender mejor las posibles adaptaciones específicas del fútbol sobre el rango de movimiento de las articulaciones de las extremidades inferiores y a individualizar el entrenamiento de la estabilidad dinámica de acuerdo a las diferencias relacionadas con el sexo. En general, esta tesis proporciona información útil para facilitar el proceso de toma de decisiones a la hora de elegir las mejores estrategias para la prevención de lesiones.

A continuación se resumen las principales contribuciones de esta tesis:

Estudio 1:

1. Los jugadores profesionales de fútbol masculino están expuestos a un riesgo sustancial de sufrir lesiones, especialmente durante los partidos.
2. La extremidad inferior fue el área corporal más lesionada, siendo el muslo la región anatómica donde las lesiones ocurrieron más habitualmente.
3. El tipo de lesión más común fueron los desgarros musculares y del tendón.
4. La mayoría de las lesiones mostraron ser de gravedad mínima y estar causadas por un mecanismo traumático.

Estudio 2:

5. El análisis llevado a cabo indicó que un gran número de jugadores de fútbol muestran valores limitados de rango de movimiento de flexión pasiva de cadera con rodilla extendida y/o dorsiflexión del tobillo con rodilla flexionada.
6. No hubo diferencias significativas en las medidas del rango de movimiento para la cadera, la rodilla y el tobillo entre jugadores de campo y porteros.
7. Se identificaron algunas diferencias bilaterales para abducción pasiva de cadera, rotación interna y externa pasiva de cadera en jugadores de fútbol profesional.

Estudio 3:

8. El modelo de predicción de lesiones musculares demostró una gran precisión para identificar a jugadores profesionales de fútbol y balonmano con alto o bajo riesgo de lesión durante las valoraciones de pre-temporada. El modelo mostró una puntuación del área bajo la curva de 0.747, con tasas de verdaderos positivos y negativos de 65.9% y 79.1%, respectivamente. La devaluación deportiva, un historial de lesión muscular en la temporada anterior y el ángulo de pico máximo de fuerza medido a través de movimientos de extensión de rodilla concéntricos (cuádriceps) y excéntricos (isquiosurales) fueron identificados como los parámetros más influyentes en la lesión muscular.

Estudio 4:

9. El modelo de predicción de lesiones isquiosurales demostró una gran precisión para identificar a jugadores profesionales de fútbol y balonmano con riesgo de lesión en los isquiosurales durante las valoraciones de pre-temporada. El modelo mostró una puntuación del área bajo la curva de 0.867, con tasas de falsos positivos y negativos de 10% y 28%, respectivamente. La calidad del sueño, la disminución del sentido de logro, un historial de lesión isquiosural en la temporada anterior, los ratios funcionales cuádriceps-isquiosurales y los valores máximos de cuádriceps-isquiosurales obtenidos cerca de la extensión de la rodilla fueron identificados como los parámetros más influyentes en la lesión de isquiosurales.

Estudio 5:

10. Los jugadores de fútbol profesional masculino y femenino estuvieron influenciados por diferentes variables de rendimiento neuromuscular en su capacidad para mantener el equilibrio dinámico unilateral.
11. La estabilidad de los futbolistas estuvo determinada por patrones de movimiento en el plano sagital, tales como el rango de movimiento de la flexión pasiva de cadera con rodilla flexionada y dorsiflexión del tobillo con rodilla flexionada.
12. La estabilidad de las futbolistas estuvo determinada por patrones de movimiento en el plano frontal, tales como la estabilidad del tronco (desplazamientos medio-laterales), la fuerza de abducción de cadera y el rango de movimiento de la dorsiflexión de tobillo con rodilla flexionada.

9.2. Limitaciones de la tesis y futuras líneas de investigación

La mayoría de las limitaciones de esta tesis doctoral han sido tratadas en la discusión de cada uno de los cinco estudios (capítulos 3, 4, 5, 6, 7). Además, esta sección presenta varias limitaciones que han sido el origen de nuevos proyectos de investigación en el Laboratorio de Biomecánica y Salud del Centro de Investigación del Deporte de la Universidad Miguel Hernández de Elche. Resumidamente, los nuevos objetivos de investigación son los siguientes:

1. *Recopilar y analizar la incidencia de lesiones en diferentes deportes de equipo (por ejemplo, baloncesto, balonmano, fútbol sala) y niveles de juego (profesionales versus aficionados).* En esta tesis solo hemos analizado la epidemiología de lesiones en fútbol masculino profesional, lo que nos ha permitido conocer las principales características de las lesiones de los jugadores y con ello establecer modelos de predicción de las lesiones con mayor tasa de incidencia. Sin embargo, no podemos afirmar que estos datos sean los mismos en otros grupos de edad, niveles de juego o sexo, y especialmente en otros deportes de equipo. Por ello, se hace indispensable que futuros trabajos investiguen estos aspectos para poder desarrollar modelos preventivos acordes a cada población o disciplina deportiva.

2. *Evaluar la aplicabilidad de los protocolos desarrollados en esta tesis doctoral a diferentes poblaciones.* Debido a las características intrínsecas del fútbol, es decir, tipo y localización de las lesiones con mayor incidencia, factores de riesgo asociados, requerimientos físicos, etc., es posible que los resultados obtenidos en los estudios tres y cuatro no puedan generalizarse a otras poblaciones. Por ello, las futuras líneas de trabajo deben explorar la aplicabilidad de estos protocolos a otros deportes de alto rendimiento con alta incidencia de lesiones, así como a personas no entrenadas, diferentes grupos de edad, niveles de juego y sexo. En este caso, una tesis doctoral que se está desarrollando en nuestro grupo de investigación está intentando replicar estos resultados en otro deporte de alto rendimiento en ambos sexos.

3. *Incluir más factores de riesgo, basados en evidencias científicas, en los modelos de predicción.* Aunque nuestros modelos de predicción de lesiones han llegado analizar más de 200 variables relacionadas con factores de riesgo descritos en la literatura, no fue posible incluir otros factores de riesgo por limitaciones metodológicas e instrumentales, que podrían llegar a mejorar los modelos obtenidos. Entre estos factores de riesgo podríamos destacar factores genéticos,[69,70] factores de la condición física como el VO_2max[79] o cinemáticos como los saltos y caidas.[321] En este sentido, algunos trabajos de nuestro grupo de investigación han incluido estos factores de riesgo en sus registros.

4. *Mejorar los enfoques estadísticos complejos procedentes de los entornos de aprendizaje automático y de minería de datos usados en esta tesis*. Los modelos matemáticos presentados en los estudios tres y cuatro siguen teniendo algunas limitaciones, porque tener un modelo con buena precisión predictiva no es suficiente si estamos interesados en responder por qué ocurrió la lesión y qué predictores están asociados con ella.[127,325] No podemos responder a estas preguntas si usamos el mismo enfoque que las técnicas de muestreo como los árboles de decisión, puesto que únicamente nos permiten dicotomizar al deportista. Por ejemplo, podríamos estar interesados en conocer si la probabilidad de lesión se modificaría si aumentase el desequilibrio de fuerza de isquiosurales entre piernas del deportista o si presentara un déficit en el rango de movimiento, lo que se podría estimar a partir de modelos estadísticos como las redes Bayesianas.

5. *Realizar test de campo que permitan crear modelos de predicción más accesibles en el campo de las ciencias del deporte.* Una de las principales limitaciones de esta tesis es la complejidad técnica de los modelos matemáticos utilizados en algunos estudios, así como la especificidad de los protocolos de laboratorio utilizados. Si bien estos atributos le proporcionan un gran valor metodológico a la tesis doctoral, hacen muy difícil replicar las principales conclusiones obtenidas. Por ello, en la actualidad, nuestro grupo de investigación está trabajando en una línea de investigación enfocada en la utilización de test de campo, menos costosos, más específicos y más fáciles de llevar cabo, con la intención de crear modelos de predicción de lesiones más próximos a la realidad del deporte de alto rendimiento. Nuestro grupo está además creando una aplicación móvil que podría ser usada en el contexto clínico, deportivo y de la investigación para predecir el riesgo de lesión, porque su uso no requiere de un amplio conocimiento de las complejas técnicas estadísticas.

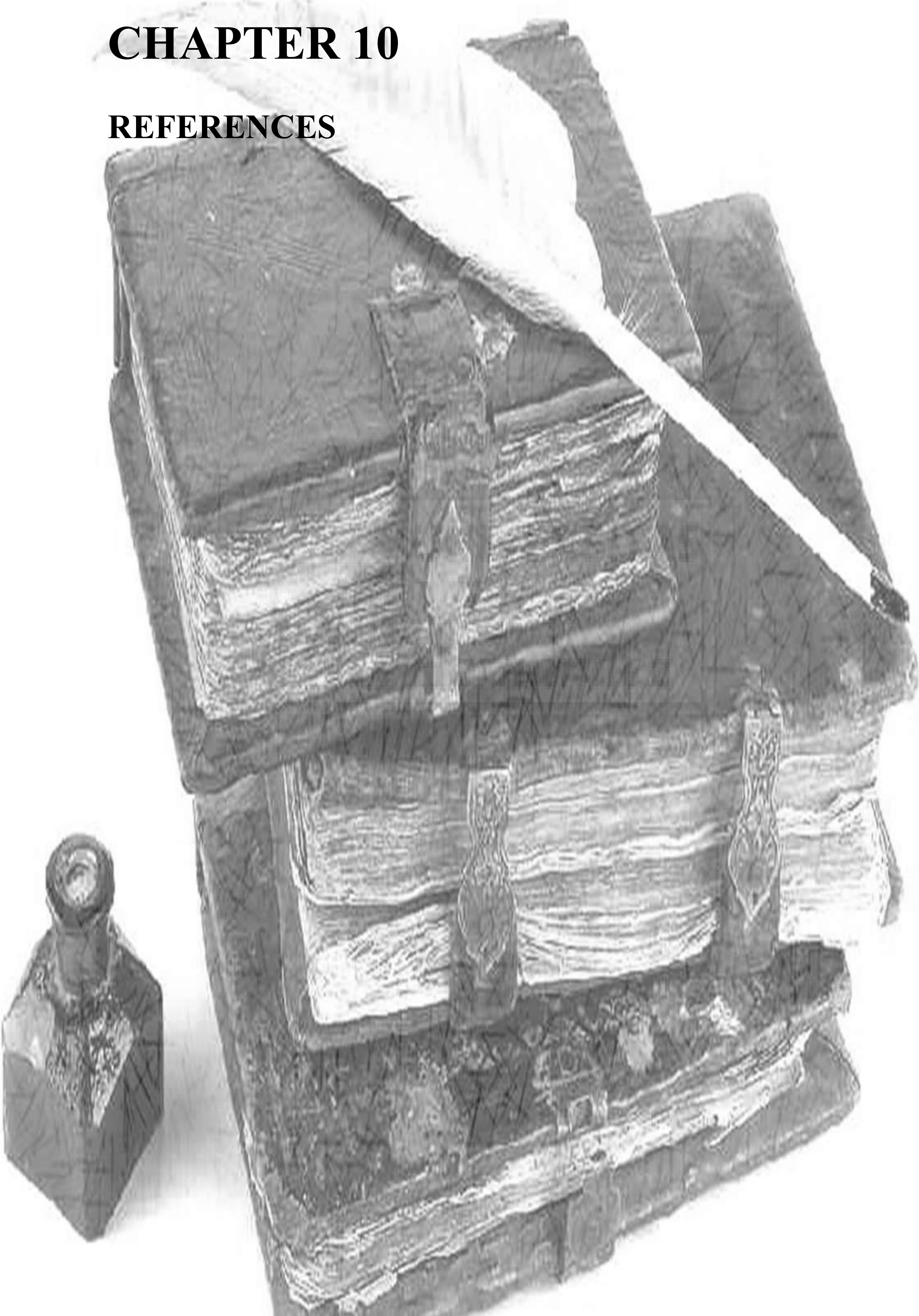
6. *Examinar posibles interacciones entre más factores de riesgo.* Esta tesis ha mostrado la interacción de varios parámetros neuromusculares (ej. rango de movimiento de la extremidad inferior y estabilidad del tronco) sobre la estabilidad unilateral dinámica. Sin embargo, muy poco parámetros fueron incluidos en los modelos de regresión, y pocos estudios han investigado las interacciones entre otros parámetros neuromusculares considerados como factores de riesgo en la literatura. Próximas investigaciones deberían investigar las relaciones entre otros parámetros neuromusculares, y/o incluir modelos estadísticos no lineales más complejos que puedan mostrar otro tipo de interacciones entre factores de riesgo, que no sean simplemente modelos-lineales.

7. *Comprobar la existencia de adaptaciones biomecánicas específicas del deporte en otras poblaciones.* A pesar de las diferentes demandas físicas y fisiológicas del fútbol, esta tesis no encontró adaptaciones específicas de los futbolistas en el rango de movimiento de la extremidad inferior respecto a los valores publicados en población sedentaria. Sin embargo, estudios en otros deportes, como balonmano, golf o tenis, si han encontrado adaptaciones específicas en el rango de movimiento de los hombros respecto a poblaciones control,[250,324,325] mientras que otros tres estudios encontraron diferencias en los valores de rango de movimiento de la cadera en jugadores de balonmano y futbol sala respecto a población no entrenada.[224,238,324] Futuras investigaciones deberían confirmar estos resultados en otros equipos profesionales y en otros niveles de juego antes de afirmar que existen o no dichas adaptaciones específicas en el rango de movimiento de la extremidad inferior en el fútbol.

8. *Desarrollar protocolos de prevención una vez los jugadores de alto riesgo han sido identificados.* Los modelos de prevención de lesiones establecen que una vez obtenida la información sobre el problema (lesión) y las principales causas del mismo, el siguiente paso es desarrollar programas preventivos que ratifiquen los datos obtenidos y se muestre la eficacia de las estrategias preventivas. Por tanto, otra de las principales limitaciones de la tesis, y línea futura de trabajo, es la necesidad de corroborar de manera práctica los principales resultados de esta tesis, estableciendo programas preventivos totalmente individualizados en función de la identificación de los deportistas con más riesgo de lesión y de los factores de riesgo asociados a cada uno.

CHAPTER 10

REFERENCES



CHAPTER 10

REFERENCES

1. FIFA C. FIFA Big Count 2006: 270 million people active in football. *FIFA Communications Division, Information Services.* 2007;31:2007.
2. Ekstrand J, Dvorak J, D'Hooghe M. Sport medicine research needs funding: the International football federations are leading the way. *British Journal of Sports Medicine*. 2013;47:726-728.
3. Carling C, Le Gall F, McCall A, Nédélec M, Dupont G. Squad management, injury and match performance in a professional soccer team over a championship-winning season. *European Journal of Sport Science.* 2015;15(7):573-582.
4. Eirale C, Tol J, Farooq A, Smiley F, Chalabi H. Low injury rate strongly correlates with team success in Qatari professional football. *British Journal of Sports Medicine.* 2013c;47(12):807-808.
5. Hägglund M, Waldén M, Magnusson H, Kristenson K, Bengtsson H, Ekstrand J. Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine.* 2013;47(12):738-742.
6. Van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries. *Sports Medicine.* 1992;14(2):82-99.
7. Bahr R. Why screening tests to predict injury do not work—and probably never will…: a critical review. *British Journal of Sports Medicine.* 2016;50:776-780.
8. Bittencourt NFN, Meeuwisse WH, Mendonça LD, et al. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. *British Journal of Sports Medicine*. 2016;50:1309-1314
9. Meeuwisse WH, Tyreman H, Hagel B, Emery C. A dynamic model of etiology in sport injury: the recursive nature of risk and causation. *Clinical Journal of Sport Medicine.* 2007;17(3):215-219.
10. Finch C. A new framework for research leading to sports injury prevention. *Journal of Science and Medicine in Sport.* 2006;9(1):3-9.
11. Van Tiggelen D, Wickes S, Stevens V, Roosen P, Witvrouw E. Effective prevention of sports injuries: a model integrating efficacy, efficiency, compliance and risk-taking behaviour. *British Journal of Sports Medicine.* 2008;42(8):648-652.
12. Bradley PS, Di Mascio M, Peart D, Olsen P, Sheldon B. High-intensity activity profiles of elite soccer players at different performance levels. *The Journal of Strength & Conditioning Research.* 2010;24(9):2343-2351.
13. Di Salvo V, Gregson W, Atkinson G, Tordoff P, Drust B. Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine.* 2009;30(03):205-212.
14. Bangsbo J. Physiological demands of football. *Sports Science Exchange.* 2014;27(125):1-6.
15. Barnes C, Archer D, Hogg B, Bush M, Bradley P. The evolution of physical and technical performance parameters in the English Premier League. *International Journal of Sports Medicine.* 2014;35(13):1095-1100.

16. Ekstrand J, Waldén M, Hägglund M. Risk for injury when playing in a national football team. *Scandinavian Journal of Medicine & Science in Sports.* 2004;14(1):34-38.
17. Bengtsson H, Ekstrand J, Hägglund M. Muscle injury rates in professional football increase with fixture congestion: an 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine.* 2013;47(12):743-747.
18. Ekstrand J, Hägglund M, Waldén M. Injury incidence and injury patterns in professional football: the UEFA injury study. *British Journal of Sports Medicine.* 2010;45:553-558.
19. Fuller CW, Ekstrand J, Junge A, et al. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Scandinavian Journal of Medicine & Science in Sports.* 2006;16(2):83-92.
20. Arnason A, Sigurdsson S, Gudmundsson A, Holme I, Engebretsen L, Bahr R. Risk factors for injuries in football. *The American Journal of Sports Medicine.* 2004;32(1 Suppl):5S-16S.
21. Hägglund M, Waldén M, Ekstrand J. Exposure and injury risk in Swedish elite football: a comparison between seasons 1982 and 2001. *Scandinavian Journal of Medicine & Science in Sports.* 2003;13(6):364-370.
22. Waldén M, Hägglund M, Ekstrand J. UEFA Champions League study: a prospective study of injuries in professional football during the 2001–2002 season. *British Journal of Sports Medicine.* 2005;39(8):542-546.
23. Hägglund M, Waldén M, Ekstrand J. Previous injury as a risk factor for injury in elite football: a prospective study over two consecutive seasons. *British Journal of Sports Medicine.* 2006;40(9):767-772.
24. Hawkins RD, Fuller CW. A prospective epidemiological study of injuries in four English professional football clubs. *British Journal of Sports Medicine.* 1999;33(3):196-203.
25. Hawkins RD, Hulse M, Wilkinson C, Hodson A, Gibson M. The association football medical research programme: an audit of injuries in professional football. *British Journal of Sports Medicine.* 2001;35(1):43-47.
26. Junge A, Dvořák J. Football injuries during the 2014 FIFA World Cup. *British Journal of Sports Medicine.* 2015;49(9):599-602.
27. Junge A, Dvorak J, Graf-Baumann T. Football injuries during the World Cup 2002. *The American Journal of Sports Medicine.* 2004b;32(1_suppl):23S-27S.
28. Junge A, Dvorak J, Graf-Baumann T, Peterson L. Football injuries during FIFA tournaments and the olympic games, 1998-2001. *The American Journal of Sports Medicine.* 2004a;32(1_suppl):80S-89S.
29. Waldén M, Hägglund M, Ekstrand J. Football injuries during European championships 2004–2005. *Knee Surgery, Sports Traumatology, Arthroscopy.* 2007;15(9):1155-1162.
30. Williams S, Trewartha G, Kemp S, Stokes K. A meta-analysis of injuries in senior men's professional Rugby Union. *Sports Medicine.* 2013;43(10):1043-1055.
31. Podlog L, Buhler CF, Pollack H, Hopkins PN, Burgess PR. Time trends for injuries and illness, and their relation to performance in the National Basketball Association. *Journal of Science and Medicine in Sport.* 2015;18(3):278-282.

32. Moller M, Attermann J, Myklebust G, Wedderkopp N. Injury risk in Danish youth and senior elite handball using a new SMS text messages approach. *British Journal of Sports Medicine.* 2012;46(7):531-537.

33. Bjørneboe J, Bahr R, Andersen TE. Risk of injury on third-generation artificial turf in Norwegian professional football. *British Journal of Sports Medicine.* 2010;44(11):794-798.

34. Hassabi M, Mortazavi SMJ, Giti MR, Hassabi M, Mansournia MA, Shapouran S. Injury profile of a professional soccer team in the premier league of iran. *Asian Journal of Sports Medicine.* 2010;1(4):201.

35. Hägglund M, Waldén M, Ekstrand J. Injury recurrence is lower at the highest professional football level than at national and amateur levels: does Sports Medicine and sports physiotherapy deliver? *British Journal of Sports Medicine.* 2016;50:751-758.

36. Mallo J, González P, Veiga S, Navarro E. Injury incidence in a Spanish sub-elite professional football team: A prospective study during four consecutive seasons. *Journal of Sports Science & Medicine.* 2011;10(4):731.

37. Hägglund M, Waldén M, Ekstrand J. UEFA injury study–an injury audit at European Championships 2006 to 2008. *British Journal of Sports Medicine.* 2009b;43:483-489.

38. Hägglund M, Waldén M, Ekstrand J. Injury incidence and distribution in elite football—a prospective study of the Danish and the Swedish top divisions. *Scandinavian Journal of Medicine & Science in Sports.* 2005;15(1):21-28.

39. Engström B, Forssblad M, Johansson C, Tornkvist H. Does a major knee injury definitely sideline an elite soccer player? *The American Journal of Sports Medicine.* 1990;18(1):101-105.

40. Gouttebarge V, Schwab BAH, Vivian A, Kerkhoffs GM. Injuries, matches missed and the influence of minimum medical standards in the A-League professional football: a 5-year prospective study. *Asian Journal of Sports Medicine.* 2016;7(1).

41. Ekstrand J, Hägglund M, Waldén M. Epidemiology of muscle injuries in professional football (soccer). *The American Journal of Sports Medicine.* 2011b;39(6):1226-1232.

42. Almutawa M, Scott M, George K, Drust B. The incidence and nature of injuries sustained on grass and 3rd generation artificial turf: a pilot study in elite Saudi National Team footballers. *Physical Therapy in Sport.* 2014;15(1):47-52.

43. Bjørneboe J, Bahr R, Andersen T. Gradual increase in the risk of match injury in Norwegian male professional football: A 6-year prospective study. *Scandinavian Journal of Medicine & Science in Sports.* 2014;24(1):189-196.

44. Wong P, Hong Y. Soccer injury in the lower extremities. *British Journal of Sports Medicine.* 2005;39(8):473-482.

45. Fuller CW, Dick RW, Corlette J, Schmalz R. Comparison of the incidence, nature and cause of injuries sustained on grass and new generation artificial turf by male and female football players. Part 1: match injuries. *British Journal of Sports Medicine.* 2007;41(suppl 1):i20-i26.

46. Fuller CW, Dick RW, Corlette J, Schmalz R. Comparison of the incidence, nature and cause of injuries sustained on grass and new generation artificial turf by male and female football players. Part 2: training injuries. *British Journal of Sports Medicine.* 2007;41(suppl 1):i27-i32.

47. English Premier League Injury Analysis: 2016/17 Season. 2017, May 24. Accessed from http://www.physioroom.com/info/english-premier-league-injury-analysis-201617-season/.

48. Ekstrand J, Waldén M, Hägglund M. Hamstring injuries have increased by 4% annually in men's professional football, since 2001: a 13-year longitudinal analysis of the UEFA Elite Club injury study. *British Journal of Sports Medicine.* 2016;50(12):731-737.

49. Hägglund M, Waldén M, Ekstrand J. Injuries among male and female elite football players. *Scandinavian Journal of Medicine & Science in Sports.* 2009a;19(6):819-827.

50. Kristenson K, Bjørneboe J, Waldén M, Andersen TE, Ekstrand J, Hägglund M. The Nordic Football Injury Audit: higher injury rates for professional football clubs with third-generation artificial turf at their home venue. *British Journal of Sports Medicine.* 2013;47(12):775-781.

51. Bradley PS, Portas MD. The relationship between preseason range of motion and muscle strain injury in elite soccer players. *The Journal of Strength & Conditioning Research.* 2007;21(4):1155-1159.

52. Daneshjoo A, Rahnama N, Mokhtar AH, Yusof A. Bilateral and unilateral asymmetries of isokinetic strength and flexibility in male young professional soccer players. *Journal of Human Kinetics.* 2013;36(1):45-53.

53. Henderson G, Barnes CA, Portas MD. Factors associated with increased propensity for hamstring injury in English Premier League soccer players. *Journal of Science and Medicine in Sport.* 2010;13(4):397-402.

54. Mosler AB, Crossley KM, Thorborg K, et al. Hip strength and range of motion: normal values from a professional football league. *Journal of Science and Medicine in Sport.* 2017;20(4):339-343.

55. Rey E, Lago-Peñas C, Casáis L, Lago-Ballesteros J. The effect of immediate post-training active and passive recovery interventions on anaerobic performance and lower limb flexibility in professional soccer players. *Journal of Human Kinetics.* 2012;31:121-129.

56. Witvrouw E, Danneels L, Asselman P, D'Have T, Cambier D. Muscle flexibility as a risk factor for developing muscle injuries in male professional soccer players a prospective study. *The American Journal of Sports Medicine.* 2003;31(1):41-46.

57. Engebretsen A, Myklebust G, Holme I, Engebretsen L, Bahr R. Intrinsic risk factors for hamstring injuries among male soccer players a prospective cohort study. *The American Journal of Sports Medicine.* 2010;38(6):1147-1153.

58. Engebretsen A, Myklebust G, Holme I, Engebretsen L, Bahr R. Intrinsic risk factors for acute ankle injuries among male soccer players: a prospective cohort study. *Scandinavian Journal of Medicine & Science in Sports.* 2010;20(3):403-410.

59. Woods C, Hawkins R, Hulse M, Hodson A. The Football Association Medical Research Programme: an audit of injuries in professional football—analysis of preseason injuries. *British Journal of Sports Medicine.* 2002;36(6):436-441.

60. Volpi P, Melegati G, Tornese D, Bandi M. Muscle strains in soccer: a five-year survey of an Italian major league team. *Knee Surgery, Sports Traumatology, Arthroscopy.* 2004;12(5):482-485.

61. Waldén M, Hägglund M, Ekstrand J. Injuries in Swedish elite football—a prospective study on injury definitions, risk for injury and injury pattern during 2001. *Scandinavian Journal of Medicine & Science in Sports.* 2005;15(2):118-125.
62. Ekstrand J, Gillquist J. Soccer injuries and their mechanisms: a prospective study. *Medicine and Science in Sports and Exercise.* 1982;15(3):267-270.
63. Askling CM, Malliaropoulos N, Karlsson J. High-speed running type or stretching-type of hamstring injuries makes a difference to treatment and prognosis. *British Journal of Sports Medicine*. 2012;46:86-87.
64. Hägglund M, Waldén M, Ekstrand J. Risk factors for lower extremity muscle injury in professional soccer: the UEFA Injury Study. *The American Journal of Sports Medicine.* 2013;41(2):327-335.
65. Ryynänen J, Dvorak J, Peterson L, et al. Increased risk of injury following red and yellow cards, injuries and goals in FIFA World Cups. *British Journal of Sports Medicine.* 2013;47(15):970-973.
66. Ryynänen J, Junge A, Dvorak J, Peterson L, Karlsson J, Börjesson M. The effect of changes in the score on injury incidence during three FIFA World Cups. *British Journal of Sports Medicine.* 2013;47:960-964.
67. Ryynänen J, Junge A, Dvorak J, et al. Foul play is associated with injury incidence: an epidemiological study of three FIFA World Cups (2002–2010). *British Journal of Sports Medicine.* 2013;47(15):986-991.
68. Bengtsson H, Ekstrand J, Waldén M, Hägglund M. Match injury rates in professional soccer vary with match result, match venue, and type of competition. *The American Journal of Sports Medicine.* 2013;41(7):1505-1510.
69. Pruna R, Artells R, Ribas J, et al. Single nucleotide polymorphisms associated with non-contact soft tissue injuries in elite professional soccer players: influence on degree of injury and recovery time. *BMC Musculoskeletal Disorders.* 2013;14(1):221.
70. Pruna R, Ribas J, Montoro JB, Artells R. The impact of single nucleotide polymorphisms on patterns of non-contact musculoskeletal soft tissue injuries in a football player population according to ethnicity. *Medicina Clinica.* 2015;144(3):105-110.
71. Dupont G, Nedelec M, McCall A, McCormack D, Berthoin S, Wisløff U. Effect of 2 soccer matches in a week on physical performance and injury rate. *The American Journal of Sports Medicine.* 2010;38(9):1752-1758.
72. Hägglund M, Zwerver J, Ekstrand J. Epidemiology of patellar tendinopathy in elite male soccer players. *The American Journal of Sports Medicine.* 2011;39(9):1906-1911.
73. Ivarsson A, Johnson U, Podlog L. Psychological predictors of injury occurrence: a prospective investigation of professional Swedish soccer players. *Journal of Sport Rehabilitation.* 2013;22(1):19-26.
74. Devantier C. Psychological predictors of injury among professional soccer players. *Sport Science Review.* 2011;20(5-6):5-36.
75. Kristenson K. *Risk factors for injury in men´ s professional football*, Linköping University Electronic Press; 2015.

76. Aoki H, O'Hata N, Kohno T, Morikawa T, Seki J. A 15-year prospective epidemiological account of acute traumatic injuries during official professional soccer league matches in Japan. *The American Journal of Sports Medicine.* 2012;40(5):1006-1014.

77. Waldén M, Hägglund M, Orchard J, Kristenson K, Ekstrand J. Regional differences in injury incidence in European professional football. *Scandinavian Journal of Medicine & Science in Sports.* 2013;23(4):424-430.

78. Lanzetti R, Ciompi A, Lupariello D, Guzzini M, De Carli A, Ferretti A. Safety of third-generation artificial turf in male elite professional soccer players in Italian major league. *Scandinavian Journal of Medicine & Science in Sports.* 2017;27(4):435-439.

79. Eriksson L, Jorfeldt L, Ekstrand J. Overuse and distorsion soccer injuries related to the player's estimated maximal aerobic work capacity. *International Journal of Sports Medicine.* 1986;7(04):214-216.

80. Malone S, Owen A, Mendes B, Hughes B, Collins K, Gabbett TJ. High-speed running and sprinting as an injury risk factor in soccer: Can well-developed physical qualities reduce the risk? *Journal of Science and Medicine in Sport.* 2017.

81. Soligard T, Nilstad A, Steffen K, et al. Compliance with a comprehensive warm-up programme to prevent injuries in youth football. *British Journal of Sports Medicine.* 2010;44(11):787-793.

82. Dvorak J, Junge A, Chomiak J, et al. Risk factor analysis for injuries in football players. *The American Journal of Sports Medicine.* 2000;28(5_suppl):69-74.

83. Ekstrand J, Hägglund M, Fuller C. Comparison of injuries sustained on artificial turf and grass by male and female elite football players. *Scandinavian Journal of Medicine & Science in Sports.* 2011a;21(6):824-832.

84. Nilstad A, Andersen TE, Bahr R, Holme I, Steffen K. Risk factors for lower extremity injuries in elite female soccer players. *The American Journal of Sports Medicine.* 2014;42(4):940-948.

85. Surve I, Schwellnus MP, Noakes T, Lombard C. A fivefold reduction in the incidence of recurrent ankle sprains in soccer players using the Sport-Stirrup orthosis. *The American Journal of Sports Medicine.* 1994;22(5):601-606.

86. Lambson RB, Barnhill BS, Higgins RW. Football cleat design and its effect on anterior cruciate ligament injuries: a three-year prospective study. *The American Journal of Sports Medicine.* 1996;24(2):155-159.

87. Arnason A, Tenga A, Engebretsen L, Bahr R. A prospective video-based analysis of injury situations in elite male football. *The American Journal of Sports Medicine.* 2004;32(6):1459-1465.

88. Eirale C, Farooq A, Smiley FA, Tol JL, Chalabi H. Epidemiology of football injuries in Asia: a prospective study in Qatar. *Journal of Science and Medicine in Sport.* 2013a;16(2):113-117.

89. Faude O, Meyer T, Federspiel B, Kindermann W. Injuries in elite German football-A media-based analysis. *Deutsche Zeitschrift fur Sportmedizin.* 2009;60(6):139-144.

90. Svensson K, Eckerman M, Alricsson M, Magounakis T, Werner S. Muscle injuries of the dominant or non-dominant leg in male football players at elite level. *Knee Surgery, Sports Traumatology, Arthroscopy.* 2016:1-5.

91. Brophy R, Silvers HJ, Gonzales T, Mandelbaum BR. Gender influences: the role of leg dominance in ACL injury among soccer players. *British Journal of Sports Medicine.* 2010;44:694-697.
92. Worrell TW. Factors associated with hamstring injuries. *Sports Medicine.* 1994;17(5):338-345.
93. Fousekis K, Tsepis E, Poulmedis P, Athanasopoulos S, Vagenas G. Intrinsic risk factors of non-contact quadriceps and hamstring strains in soccer: a prospective study of 100 professional players. *British Journal of Sports Medicine.* 2011;45:709-714.
94. Shelbourne KD, Kerr B. The relationship of femoral intercondylar notch width to height, weight, and sex in patients with intact anterior cruciate ligaments. *The American Journal of Knee Surgery.* 2000;14(2):92-96.
95. Uhorchak JM, Scoville CR, Williams GN, Arciero RA, Pierre PS, Taylor DC. Risk factors associated with noncontact injury of the anterior cruciate ligament. *The American Journal of Sports Medicine.* 2003;31(6):831-842.
96. LaPrade RF, Burnett QM. Femoral intercondylar notch stenosis and correlation to anterior cruciate ligament injuries: a prospective study. *The American Journal of Sports Medicine.* 1994;22(2):198-203.
97. Butler RJ, Lehr ME, Fink ML, Kiesel KB, Plisky PJ. Dynamic balance performance and noncontact lower extremity injury in college football players: an initial study. *Sports health.* 2013;5(5):417-422.
98. Plisky PJ, Rauh MJ, Kaminski TW, Underwood FB. Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. *Journal of Orthopaedic & Sports Physical Therapy.* 2006;36(12):911-919.
99. Nielsen AB, Yde J. Epidemiology and traumatology of injuries in soccer. *The American Journal of Sports Medicine.* 1989;17(6):803-807.
100. Ekstrand J, Tropp H. The incidence of ankle sprains in soccer. *Foot & Ankle.* 1990;11(1):41-44.
101. Hawkins RD, Fuller CW. An examination of the frequency and severity of injuries and incidents at three levels of professional football. *British Journal of Sports Medicine.* 1998;32(4):326-332.
102. Inklaar H. Soccer injuries. *Sports Medicine.* 1994;18(1):55-73.
103. Azubuike S, Okojie O. An epidemiological study of football (soccer) injuries in Benin City, Nigeria. *British Journal of Sports Medicine.* 2009;43(5):382-386.
104. Peterson L, Junge A, Chomiak J, Graf-Baumann T, Dvorak J. Incidence of football injuries and complaints in different age groups and skill-level groups. *The American Journal of Sports Medicine.* 2000;28(5_suppl):51-57.
105. Ekstrand J, Hägglund M, Waldén M. Injury incidence and injury patterns in professional football: the UEFA injury study. *British Journal of Sports Medicine.* 2010;45:553-558.
106. Dvorak J, Junge A, Grimm K, Kirkendall D. Medical report from the 2006 FIFA world cup Germany. *British Journal of Sports Medicine.* 2007;41(9):578-581.
107. Rahnama N, Reilly T, Lees A. Injury risk associated with playing actions during competitive soccer. *British Journal of Sports Medicine.* 2002;36(5):354-359.
108. Laux P, Krumm B, Diers M, Flor H. Recovery–stress balance and injury risk in professional football players: a prospective study. *Journal of Sports Sciences.* 2015;33(20):2140-2148.

109. Luke A, Lazaro RM, Bergeron MF, et al. Sports-related injuries in youth athletes: is overscheduling a risk factor? *Clinical Journal of Sport Medicine.* 2011;21(4):307-314.

110. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes. *The American Journal of Sports Medicine.* 2005;33(4):492-501.

111. Smith HC, Johnson RJ, Shultz SJ, et al. A prospective evaluation of the Landing Error Scoring System (LESS) as a screening tool for anterior cruciate ligament injury risk. *The American Journal of Sports Medicine.* 2012;40(3):521-526.

112. Krosshaug T, Steffen K, Kristianslund E, et al. The vertical drop jump is a poor screening test for ACL injuries in female elite soccer and handball players: a prospective cohort study of 710 athletes. *The American Journal of Sports Medicine.* 2016;44(4):874-883.

113. Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).* 2012;42(4):463-484.

114. López V, Fernández A, García S, Palade V, Herrera F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences.* 2013;250:113-141.

115. Jovanovic M. Uncertainty, heuristics and injury prediction. *Aspetar Sports Medicine Journal.* 2017;6:18-24.

116. Bourne MN, Opar DA, Williams MD, Shield AJ. Eccentric knee flexor strength and risk of hamstring injuries in rugby union: a prospective study. *The American Journal of Sports Medicine.* 2015;43(11):2663-2670.

117. Brockett CL, Morgan DL, Proske U. Predicting hamstring strain injury in elite athletes. *Medicine & Science in Sports & Exercise.* 2004;36(3):379-387.

118. Croisier JL, Ganteaume S, Binet J, Genty M, Ferret JM. Strength imbalances and prevention of hamstring injury in professional soccer players a prospective study. *The American Journal of Sports Medicine.* 2008;36(8):1469-1475.

119. Croisier JL, Forthomme B, Namurois MH, Vanderthommen M, Crielaard JM. Hamstring muscle strain recurrence and strength performance disorders. *The American Journal of Sports Medicine.* 2002;30(2):199-203.

120. Dauty M, Menu P, Fouasson-Chailloux A, Ferréol S, Dubois C. Prediction of hamstring injury in professional soccer players by isokinetic measurements. *Muscles, Ligaments and Tendons Journal.* 2016;6(1):116.

121. Gabbe BJ, Bennell KL, Finch CF, Wajswelner H, Orchard J. Predictors of hamstring injury at the elite level of Australian football. *Scandinavian Journal of Medicine & Science in Sports.* 2006;16(1):7-13.

122. Opar DA, Williams MD, Timmins RG, Hickey J, Duhig SJ, Shield AJ. The effect of previous hamstring strain injuries on the change in eccentric hamstring strength during preseason training in elite Australian footballers. *The American Journal of Sports Medicine.* 2015;43(2):377-384.

123. Orchard J, Marsden J, Lord S, Garlick D. Preseason hamstring muscle weakness associated with hamstring muscle injury in Australian footballers. *The American Journal of Sports Medicine.* 1997;25(1):81-85.
124. Sugiura Y, Saito T, Sakuraba K, Sakuma K, Suzuki E. Strength deficits identified with concentric action of the hip extensors and eccentric action of the hamstrings predispose to hamstring injury in elite sprinters. *Journal of Orthopaedic & Sports Physical Therapy.* 2008;38(8):457-464.
125. van Dyk N, Bahr R, Whiteley R, et al. Hamstring and quadriceps isokinetic strength deficits are weak risk factors for hamstring strain injuries: A 4-year cohort study. *The American Journal of Sports Medicine.* 2016;44(7):1789-1795.
126. Mendiguchia J, Alentorn-Geli E, Brughelli M. Hamstring strain injuries: are we heading in the right direction? *British Journal of Sports Medicine.* 2012;46(2):81-85.
127. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning.* Vol 6: Springer; 2013.
128. Ofoghi B, Zeleznikow J, MacMahon C, Dwyer D. Supporting athlete selection and strategic planning in track cycling omnium: A statistical and machine learning approach. *Information Sciences.* 2013;233:200-213.
129. Chen I, Homma H, Jin C, Yan H. Identification of elite swimmers' race patterns using cluster analysis. *International Journal of Sports Science & Coaching.* 2007;2(3):293-303.
130. Edelmann-Nusser J, Hohmann A, Henneberg B. Modeling and prediction of competitive performance in swimming upon neural networks. *European Journal of Sport Science.* 2002;2(2):1-10.
131. Lamb P, Bartlett R, Robins A. Self-organising maps: An objective method for clustering complex human movement. *International Journal of Computer Science in Sport.* 2010;9(1):20-29.
132. Bhandari I, Colet E, Parker J, Pines Z, Pratap R, Ramanujam K. Advanced scout: Data mining and knowledge discovery in NBA data. *Data Mining and Knowledge Discovery.* 1997;1(1):121-125.
133. Kampakis S. Predictive modelling of football injuries. *arXiv preprint arXiv:1609.07480.* 2016.
134. Farmer JD, Gallegati M, Hommes C, et al. A complex systems approach to constructing better models for managing financial markets and the economy. *The European Physical Journal Special Topics.* 2012;214(1):295-324.
135. Batuwita R, Palade V. microPred: effective classification of pre-miRNAs for human miRNA gene prediction. *Bioinformatics.* 2009;25(8):989-995.
136. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural networks.* 2008;21(2):427-436.
137. Mena AJP, González MJC. Detección de Betalactamasas de Espectro Extendido en cepas de la familia Enterobacteriaceae. *Kasmera.* 2009;37(1).
138. Song Y, Ying L. Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry.* 2015;27(2):130.
139. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and systems magazine.* 2006;6(3):21-45.

140. Gorman PP, Butler RJ, Rauh MJ, Kiesel K, Plisky PJ. Differences in dynamic balance scores in one sport versus multiple sport high school athletes. *International Journal of Sports Physical Therapy.* 2012;7(2).

141. Booysen MJ, Gradidge PJL, Watson E. The relationships of eccentric strength and power with dynamic balance in male footballers. *Journal of Sports Sciences.* 2015;33(20):2157-2165.

142. Lockie RG, Schultz AB, Callaghan SJ, Jeffriess MD, Berry SP. Reliability and validity of a new test of change-of-direction speed for field-based sports: the change-of-direction and acceleration test (CODAT). *Journal of Sports Science & Medicine.* 2013;12(1):88.

143. Thorpe JL, Ebersole KT. Unilateral balance performance in female collegiate soccer athletes. *The Journal of Strength & Conditioning Research.* 2008;22(5):1429-1433.

144. Ambegaonkar JP, Mettinger LM, Caswell SV, Burtt A, Cortes N. Relationships between core endurance, hip strength, and balance in collegiate female athletes. *International Journal of Sports Physical Therapy.* 2014;9(5).

145. Ozmen T. Relationship between core stability, dynamic balance and jumping performance in soccer players. *Turkish Journal of Sport and Exercise.* 2016;18(1):110-113.

146. Overmoyer GV, Reiser RF. Relationships Between Lower-Extremity Flexibility, Asymmetries, and the Y Balance Test. *The Journal of Strength & Conditioning Research.* 2015;29(5):1240-1247.

147. Robinson R, Gribble P. Kinematic predictors of performance on the Star Excursion Balance Test. *Journal of Sport Rehabilitation.* 2008;17(4):347-357.

148. Chimera NJ, Smith CA, Warren M. Injury history, sex, and performance on the functional movement screen and Y balance test. *Journal of Athletic Training.* 2015;50(5):475-485.

149. Sabin MJ, Ebersole KT, Martindale AR, Price JW, Broglio SP. Balance performance in male and female collegiate basketball athletes: influence of testing surface. *The Journal of Strength & Conditioning Research.* 2010;24(8):2073-2078.

150. Gribble PA, Hertel J. Considerations for normalizing measures of the Star Excursion Balance Test. *Measurement in Physical Education and Exercise Science.* 2003;7(2):89-100.

151. Dellal A, Lago-Peñas C, Rey E, Chamari K, Orhant E. The effects of a congested fixture period on physical performance, technical activity and injury rate during matches in a professional soccer team. *British Journal of Sports Medicine.* 2015;49:390-394.

152. Bradley PS, Sheldon W, Wooster B, Olsen P, Boanas P, Krustrup P. High-intensity running in English FA Premier League soccer matches. *Journal of Sports Sciences.* 2009;27(2):159-168.

153. McCall A, Carling C, Davison M, et al. Injury risk factors, screening tests and preventative strategies: a systematic review of the evidence that underpins the perceptions and practices of 44 football (soccer) teams from various premier leagues. *British Journal of Sports Medicine.* 2015;49:583-589.

154. Di Salvo V, Benito P, Calderon F, Di Salvo M, Pigozzi F. Activity profile of elite goalkeepers during football match-play. *Journal of Sports Medicine and Physical Fitness.* 2008;48(4):443.

155. Di Salvo V, Baron R, Tschan H, Montero FC, Bachl N, Pigozzi F. Performance characteristics according to playing position in elite soccer. *International Journal of Sports Medicine.* 2007;28(03):222-227.

156. Chimera NJ, Swanik KA, Swanik CB, Straub SJ. Effects of plyometric training on muscle-activation strategies and performance in female athletes. *Journal of Athletic Training.* 2004;39(1):24.
157. Åman M, Forssblad M, Henriksson-Larsén K. Incidence and severity of reported acute sports injuries in 35 sports using insurance registry data. *Scandinavian Journal of Medicine & Science in Sports.* 2016;26(4):451-462.
158. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *Journal of Athletic Training.* 2007;42(2):311.
159. Ekstrand J. Keeping your top players on the pitch: the key to football medicine at a professional level. *British Journal of Sports Medicine.* 2013;47(12):723-724.
160. Hutton B, Catalá-López F, Moher D. The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA. *Medicina Clínica (English Edition).* 2016;147(6):262-266.
161. Hägglund M, Waldén M, Bahr R, Ekstrand J. Methods for epidemiological study of injuries to professional football players: developing the UEFA model. *British Journal of Sports Medicine.* 2005;39(6):340-346.
162. Doherty C, Delahunt E, Caulfield B, Hertel J, Ryan J, Bleakley C. The incidence and prevalence of ankle sprain injury: a systematic review and meta-analysis of prospective epidemiological studies. *Sports Medicine.* 2014;44(1):123-140.
163. Sanderson S, Tatt ID, Higgins J. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International Journal of Epidemiology.* 2007;36(3):666-676.
164. Castaneda L, Bergmann A, Bahia L. The International Classification of Functioning, Disability and Health: a systematic review of observational studies. *Revista Brasileira de Epidemiologia.* 2014;17(2):437-451.
165. Weiss K, Whatman C. Biomechanics associated with patellofemoral pain and ACL injuries in sports. *Sports Medicine.* 2015;45(9):1325-1337.
166. Orwin RG, Vevea JL. Evaluating coding decisions. *The handbook of research synthesis and meta-analysis.* 2009;2:177-203.
167. Lystad RP, Pollard H, Graham PL. Epidemiology of injuries in competition taekwondo: A meta-analysis of observational studies. *Journal of Science and Medicine in Sport.* 2009;12(6):614-621.
168. Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in medicine.* 2002;21(11):1539-1558.
169. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software.* 2010;36(3):1-48.
170. Marshall SW, Fincher GE. A spreadsheet for Combining Outcomes from Several Subject Groups. *Sportscience.* 2006;10:51-54.
171. Hopkins W. Linear models and effect magnitudes for research, clinical and practical applications. *Sportscience.* 2010;14(1):49-57.
172. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance.* 2006;1(1):50-57.

173. Hopkins W. A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a P value. *Sportscience.* 2007;11:16-21.

174. Arnason A, Engebretsen L, Bahr R. No effect of a video-based awareness program on the rate of soccer injuries. *The American Journal of Sports Medicine.* 2005;33(1):77-84.

175. Hawkins RD, Fuller CW. Risk assessment in professional football: an examination of accidents and incidents in the 1994 World Cup finals. *British Journal of Sports Medicine.* 1996;30(2):165-170.

176. Carling C, McCall A, Le Gall F, Dupont G. The impact of in-season national team soccer play on injury and player availability in a professional club. *Journal of Sports Sciences.* 2015;33(17):1751-1757.

177. Dvorak J, Junge A, Derman W, Schwellnus M. Injuries and illnesses of football players during the 2010 FIFA World Cup. *British Journal of Sports Medicine.* 2011;45:626-630..

178. Eirale C, Tol JL, Smiley F, Farooq A, Chalabi H. Does Ramadan affect the risk of injury in professional football? *Clinical Journal of Sport Medicine.* 2013b;23(4):261-266.

179. Eirale C, Hamilton B, Bisciotti G, Grantham J, Chalabi H. Injury epidemiology in a national football team of the Middle East. *Scandinavian Journal of Medicine & Science in Sports.* 2012;22(3):323-329.

180. Ekstrand J, Hägglund M, Kristenson K, Magnusson H, Waldén M. Fewer ligament injuries but no preventive effect on muscle injuries and severe injuries: an 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine.* 2013;47(12):732-737.

181. Fünten K, Faude O, Lensch J, Meyer T. Injury characteristics in the German professional male soccer leagues after a shortened winter break. *Journal of Athletic Training.* 2014;49(6):786-793.

182. Gustafsson T. *Injury incidence and injury patterns among male elite football players when playing in the national team.* 2011. Linnaeus University.

183. Lee JWY, Mok KM, Chan HCK, Yung PSH, Chan KM. A prospective epidemiological study of injury incidence and injury patterns in a Hong Kong male professional football league during the competitive season. *Asia-Pacific Journal of Sports Medicine, Arthroscopy, Rehabilitation and Technology.* 2014;1(4):119-125.

184. Noya J, Gómez-Carmona PM, Moliner-Urdiales D, Gracia-Marco L, Sillero-Quintana M. An examination of injuries in Spanish Professional Soccer League. *The Journal of Sports Medicine and Physical Fitness.* 2014a;54(6):765-771.

185. Noya J, Gómez-Carmona PM, Gracia-Marco L, Moliner-Urdiales D, Sillero-Quintana M. Epidemiology of injuries in First Division Spanish football. *Journal of Sports Sciences.* 2014b;32(13):1263-1270.

186. Poulsen T, Freund K, Madsen F, Sandvej K. Injuries in high-skilled and low-skilled soccer: a prospective study. *British Journal of Sports Medicine.* 1991;25(3):151-153.

187. Reis GF, Santos TR, Lasmar RC, Oliveira Júnior O, Lopes RF, Fonseca ST. Sports injuries profile of a first division Brazilian soccer team: a descriptive cohort study. *Brazilian Journal of Physical Therapy.* 2015;19(5):390-397.

188. Shalaj I, Tishukaj F, Bachl N, Tschan H, Wessner B, Csapo R. Injuries in professional male football players in Kosovo: a descriptive epidemiological study. *BMC Musculoskeletal Disorders.* 2016;17(1):338.
189. Stubbe JH, van Beijsterveldt AMM, van der Knaap S, et al. Injuries in professional male soccer players in the Netherlands: a prospective cohort study. *Journal of Athletic Training.* 2015;50(2):211-216.
190. Andersen TE, Tenga A, Engebretsen L, Bahr R. Video analysis of injuries and incidents in Norwegian professional football. *British Journal of Sports Medicine.* 2004;38(5):626-631.
191. Carling C, Orhant E, LeGall F. Match injuries in professional soccer: inter-seasonal variation and effects of competition type, match congestion and positional role. *International Journal of Sports Medicine.* 2010;31(04):271-276.
192. Junge A, Dvorak J. Injury surveillance in the world football tournaments 1998–2012. *British Journal of Sports Medicine.* 2013;47:782-788.
193. Pedrinelli A, da Cunha Filho GAR, Thiele ES, Kullak OP. Epidemiological study on professional football injuries during the 2011 Copa America, Argentina. *Revista Brasileira de Ortopedia (English Edition).* 2013;48(2):131-136.
194. Arnason A, Gudmundsson A, Dahl H, Johannsson E. Soccer injuries in Iceland. *Scandinavian Journal of Medicine & Science in Sports.* 1996;6(1):40-45.
195. Stubbe JH, Schmikli S, van de Port I, Backx FJ. Differences in injury risk and characteristics between Dutch amateur and professional soccer players. *Journal of Science and Medicine in Sport.* 2015;18(2):145-149.
196. Williams S, Trewartha G, Kemp SPT, et al. Time loss injuries compromise team success in Elite Rugby Union: a 7-year prospective study. *British Journal of Sports Medicine.* 2016;50:651-656.
197. Hägglund M. *Epidemiology and prevention of football injuries*, Institutionen för hälsa och samhälle; 2007.
198. Bangsbo J, Mohr M, Krustrup P. Physical and metabolic demands of training and match-play in the elite football player. *Journal of Sports Sciences.* 2006;24(07):665-674.
199. Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *British Journal of Sports Medicine.* 2016;50:273-280.
200. Corazza A, Orlandi D, Baldari A, et al. Thigh muscles injuries in professional soccer players: a one year longitudinal study. *Muscles, Ligaments and Tendons Journal.* 2013;3(4):331.
201. Petersen J, Thorborg K, Nielsen MB, Budtz-Jørgensen E, Hölmich P. Preventive effect of eccentric training on acute hamstring injuries in men's soccer: a cluster-randomized controlled trial. *The American Journal of Sports Medicine.* 2011;39(11):2296-2303.
202. Hennessey L, Watson A. Flexibility and posture assessment in relation to hamstring injury. *British Journal of Sports Medicine.* 1993;27(4):243-246.
203. Fousekis K, Tsepis E, Poulmedis P, Athanasopoulos S, Vagenas G. Intrinsic risk factors of non-contact quadriceps and hamstring strains in soccer: a prospective study of 100 professional players. *British Journal of Sports Medicine.* 2011;45(9):709-714.

204. Askling CM, Nilsson J, Thorstensson A. A new hamstring test to complement the common clinical examination before return to sport after injury. *Knee Surgery, Sports Traumatology, Arthroscopy.* 2010;18(12):1798-1803.

205. Lees A, Nolan L. The biomechanics of soccer: a review. *Journal of Sports Sciences.* 1998;16(3):211-234.

206. Lee M, Reid SL, Elliott BC, Lloyd DG. Running biomechanics and lower limb strength associated with prior hamstring injury. *Medicine & Science in Sports & Exercise.* 2009;41(10):1942-1951.

207. Timmins RG, Bourne MN, Shield AJ, Williams MD, Lorenzen C, Opar DA. Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. *British Journal of Sports Medicine.* 2016;50:1524-1535.

208. Rahnama N, Reilly T, Lees A, Graham-Smith P. Muscle fatigue induced by exercise simulating the work rate of competitive soccer. *Journal of Sports Science.* 2003;21(11):933-942.

209. Thorlund JB, Aagaard P, Madsen K. Rapid muscle force capacity changes after soccer match play. *International Journal of Sports Medicine.* 2009;30(04):273-278.

210. Emery C, Meeuwisse W. The effectiveness of a neuromuscular prevention strategy to reduce injuries in youth soccer: a cluster-randomised controlled trial. *British Journal of Sports Medicine.* 2010;44(8):555-562.

211. Owoeye OB, Akinbo SR, Tella BA, Olawale OA. Efficacy of the FIFA 11+ warm-up programme in male youth football: a cluster randomised controlled trial. *Journal of Sports Science & Medicine.* 2014;13(2):321.

212. Silvers H, Mandelbaum B, Bizzini M, Dvorak J. The efficacy of the FIFA 11+ program in the collegiate male soccer player (USA). *British Journal of Sports Medicine.* 2014;48(7):662-662.

213. Soligard T, Myklebust G, Steffen K, et al. Comprehensive warm-up programme to prevent injuries in young female footballers: cluster randomised controlled trial. *British Medical Journal.* 2008;337:a2469.

214. Lystad RP, Gregory K, Wilson J. The epidemiology of injuries in mixed martial arts a systematic review and meta-analysis. *Orthopaedic Journal of Sports Medicine.* 2014;2(1):2325967113518492.

215. Krustrup P, Aagaard P, Nybo L, Petersen J, Mohr M, Bangsbo J. Recreational football as a health promoting activity: a topical review. *Scandinavian Journal of Medicine & Science in Sports.* 2010;20(s1):1-13.

216. Witvrouw E, Mahieu N, Danneels L, McNair P. Stretching and injury prevention. *Sports Medicine.* 2004;34(7):443-449.

217. Rahnama N, Lees A, Bambaecichi E. A comparison of muscle strength and flexibility between the preferred and non-preferred leg in English soccer players. *Ergonomics.* 2005;48(11-14):1568-1575.

218. Wilcox CR, Osgood CT, White HS, Vince RV. Investigating strength and range of motion of the hip complex in ice hockey athletes. *Journal of Sport Rehabilitation.* 2015;24(3):300-306.

219. Manning C, Hudson Z. Comparison of hip joint range of motion in professional youth and senior team footballers with age-matched controls: an indication of early degenerative change? *Physical Therapy in Sport.* 2009;10(1):25-29.

220. Beynnon BD, Renström PA, Alosa DM, Baumhauer JF, Vacek PM. Ankle ligament injury risk factors: a prospective study of college athletes. *Journal of Orthopaedic Research.* 2001;19(2):213-220.

221. Murphy D, Connolly D, Beynnon B. Risk factors for lower extremity injury: a review of the literature. *British Journal of Sports Medicine.* 2003;37(1):13-29.

222. Drawer S, Fuller C. Propensity for osteoarthritis and lower limb joint pain in retired professional soccer players. *British Journal of Sports Medicine.* 2001;35(6):402-408.

223. Öberg B, Ekstrand J, Möller M, Gillquist J. Muscle strength and flexibility in different positions of soccer players. *International Journal of Sports Medicine.* 1984;5(04):213-216.

224. Cejudo A, Sainz de Baranda P, Ayala F, Santonja F. Perfil de flexibilidad de la extremidad inferior en jugadores de fútbol sala. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte.* 2014; 14(55).

225. Heck C, Hendryson I, Rowe C. Joint motion: method of measuring and recording. *Chicago, IL: American Academy of Orthopedic Surgeons.* 1965:30-42.

226. Association AM. *Guides to the evaluation of permanent impairment.* American Medical Association Press; 1990.

227. Magee DJ. Orthopedic physical assessment. *Philadelphia: WB Saunders.* 1997;3.

228. Prentice W. The thigh, hip, groin, and pelvis. *Arnheim's principles of athletic training: A competency-based approach.* 2003:656-698.

229. Cejudo A, Sainz de Baranda P, Ayala F, Santonja F. A simplified version of the weight-bearing ankle lunge test: Description and test–retest reliability. *Manual therapy.* 2014;19(4):355-359.

230. Taylor KL, Sheppard JM, Lee H, Plummer N. Negative effect of static stretching restored when combined with a sport specific warm-up component. *Journal of Science and Medicine in Sport.* 2009;12(6):657-661.

231. Ayala F, Moreno-Pérez V, Vera-Garcia FJ, Moya M, Sanz-Rivas D, Fernandez-Fernandez J. Acute and time-course effects of traditional and dynamic warm-up routines in young elite junior tennis players. *PloS one.* 2016;11(4):e0152790.

232. Holla JF, van der Leeden M, Roorda LD, et al. Diagnostic accuracy of range of motion measurements in early symptomatic hip and/or knee osteoarthritis. *Arthritis Care & Research.* 2012;64(1):59-65.

233. Kendall FP, McCreary EK, Provance PG, Rodgers M, Romani WA. Muscles: Testing and function, with posture and pain: Includes a bonus primal anatomy. *Philadelphia, PA: Lippincott Williams & Wilkins.* 2005;5:158.

234. Pope R, Herbert R, Kirwan J. Effects of ankle dorsiflexion range and pre-exercise calf muscle stretching on injury risk in Army recruits. *Australian Journal of Physiotherapy.* 1998;44(3):165-172.

235. Roach S, San Juan JG, Suprak DN, Lyda M. Concurrent validity of digital inclinometer and universal goniometer in assessing passive hip mobility in healthy subjects. *International Journal of Sports Physical Therapy.* 2013;8(5):680.

236. Young W, Clothier P, Otago L, Bruce L, Liddell D. Acute effects of static stretching on hip flexor and quadriceps flexibility, range of motion and foot speed in kicking a football. *Journal of Science and Medicine in Sport.* 2004;7(1):23-31.

237. Gerhardt JJ, Cocchiarella L, Lea RD. *The practical guide to range of motion assessment.* Amer Medical Assn; 2002.

238. L'hermette M, Polle G, Tourny-Chollet C, Dujardin F. Hip passive range of motion and frequency of radiographic hip osteoarthritis in former elite handball players. *British Journal of Sports Medicine.* 2006;40(1):45-49.

239. Kibler W, McQueen C, Uhl T. Fitness evaluations and fitness findings in competitive junior tennis players. *Clinics in Sports Medicine.* 1988;7(2):403-416.

240. Hopkins W. A scale of magnitudes for effect statistics. *A new view of statistics.* 2002;502.

241. Cohen J. Statistical Power Analysis for the Behavioral SciencesNew JerseyLawrence Erlbaum Associates. *Inc. Publishers.* 1988.

242. Komi PV, Bosco C. Muscles by men and women. *Medicine & Science in Sports & Exercise.* 1978;10:261-265.

243. Orchard J. Hamstrings are most susceptible to injury during the early stance phase of sprinting. *British Journal of Sports Medicine.* 2012;46(2):88-89.

244. Sun Y, Wei S, Zhong Y, Fu W, Li L, Liu Y. How joint torques affect hamstring injury risk in sprinting swing–stance transition. *Medicine and Science in Sports and Exercise.* 2015;47(2):373.

245. Friden J, Lieber R. Eccentric exercise-induced injuries to contractile and cytoskeletal muscle fibre components. *Acta Physiologica Scandinavica.* 2001;171(3):321-326.

246. Witvrouw E, Bellemans J, Lysens R, Danneels L, Cambier D. Intrinsic Risk Factors for the Development of Patellar Tendinitis in an Athletic Population A Two-Year Prospective Study. *The American Journal of Sports Medicine.* 2001;29(2):190-195.

247. Fletcher IM, Monte-Colombo MM. An investigation into the possible physiological mechanisms associated with changes in performance related to acute responses to different preactivity stretch modalities. *Applied Physiology, Nutrition, and Metabolism.* 2010;35(1):27-34.

248. Simic L, Sarabon N, Markovic G. Does pre-exercise static stretching inhibit maximal muscular performance? A meta-analytical review. *Scandinavian Journal of Medicine & Science in Sports.* 2013;23(2):131-148.

249. Moreno-Pérez V, Moreside J, Barbado D, Vera-Garcia FJ. Comparison of shoulder rotation range of motion in professional tennis players with and without history of shoulder pain. *Manual Therapy.* 2015;20(2):313-318.

250. Ibrahim A, Murrell G, Knapman P. Adductor strain and hip range of movement in male professional soccer players. *Journal of Orthopaedic Surgery.* 2007;15(1):46-49.

251. Brooks JH, Fuller CW, Kemp SP, Reddin DB. Incidence, risk, and prevention of hamstring muscle injuries in professional rugby union. *The American Journal of Sports Medicine.* 2006;34(8):1297-1306.

252. Langevoort G, Myklebust G, Dvorak J, Junge A. Handball injuries during major international tournaments. *Scandinavian Journal of Medicine & Science in Sports.* 2007;17(4):400-407.

253. Mendiguchia J, Alentorn-Geli E, Idoate F, Myer GD. Rectus femoris muscle injuries in football: a clinically relevant review of mechanisms of injury, risk factors and preventive strategies. *British Journal of Sports Medicine.* 2013;47(6):359-366.
254. Faude O, Rößler R, Junge A. Football injuries in children and adolescent players: are there clues for prevention? *Sports Medicine.* 2013;43(9):819-837.
255. Olmedilla A, Laguna M, Redondo AB. Lesiones y características psicológicas en jugadores de balonmano. *Revista Andaluza de Medicina del Deporte.* 2011;4(1):6-12.
256. Macías J, Royuela A. La versión española del Índice de Calidad de Sueño de Pittsburgh. *Informaciones Psiquiátricas.* 1996;146(4):465-472.
257. Arce C, De Francisco C, Andrade E, Seoane G, Raedeke T. Adaptation of the Athlete Burnout Questionnaire in a Spanish sample of athletes. *The Spanish Journal of Psychology.* 2012;15(03):1529-1536.
258. Shaffer SW, Teyhen DS, Lorenson CL, et al. Y-balance test: a reliability study involving multiple raters. *Military Medicine.* 2013;178(11):1264-1270.
259. Thorborg K, Petersen J, Magnusson S, Hölmich P. Clinical assessment of hip strength using a hand-held dynamometer is reliable. *Scandinavian Journal of Medicine & Science in Sports.* 2010;20(3):493-501.
260. Barbado D, Lopez-Valenciano A, Juan-Recio C, Montero-Carretero C, van Dien JH, Vera-Garcia FJ. Trunk stability, trunk strength and sport performance level in judo. *PloS one.* 2016;11(5):e0156267.
261. Hancock GR, Butler MS, Fischman MG. On the problem of two-dimensional error scores: Measures and analyses of accuracy, bias, and consistency. *Journal of Motor Behavior.* 1995;27(3):241-250.
262. Ayala F, Puerta JM, Flores MJ, et al. Análisis bayesiano de los principales factores de riesgo de lesión de la musculatura isquiosural. *Kronos.* 2016;15(1).
263. Elkarami B, Alkhateeb A, Rueda L. Cost-sensitive classification on class-balanced ensembles for imbalanced non-coding RNA data. *Paper presented at: Student Conference (ISC), 2016 IEEE EMBS International*. 2016, p. 1-4.
264. Hacibeyoglu M, Arslan A, Kahramanli S. Improving Classification Accuracy with Discretization on Data Sets Including Continuous Valued Features. *Ionosphere.* 2011;34(351):2.
265. Quinlan JR. Learning decision tree classifiers. *ACM Computing Surveys (CSUR).* 1996;28(1):71-72.
266. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Wadsworth & Brooks. *Monterey, CA.* 1984.
267. Freund Y, Mason L. The alternating decision tree learning algorithm. Paper presented at: icml1999.
268. Aldous D. The continuum random tree. I. *The Annals of Probability.* 1991:1-28.
269. Refaeilzadeh P, Tang L, Liu H. Cross-validation. *Encyclopedia of database systems*: Springer; 2009:532-538.
270. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatrica.* 2007;96(5):644-647.
271. Cresswell SL, Eklund RC. The nature of player burnout in rugby: Key characteristics and attributions. *Journal of Applied Sport Psychology.* 2006;18(3):219-239.

272. Palmer ML, Epler ME. *Fundamentos de las técnicas de evaluación musculoesquelética.* Vol 85: Editorial Paidotribo; 2002.

273. Croisier JL, Réveillon V, Ferret J, et al. Isokinetic assessment of knee flexors and extensors in professional soccer players. *Isokinetics & Exercise Science.* 2003;11(1):61-62.

274. Yeung SS, Suen AM, Yeung EW. A prospective cohort study of hamstring injuries in competitive sprinters: preseason muscle imbalance as a possible risk factor. *British Journal of Sports Medicine.* 2009;43(8):589-594.

275. Devan MR, Pescatello LS, Faghri P, Anderson J. A prospective study of overuse knee injuries among female athletes with muscle imbalances and structural abnormalities. *Journal of Athletic Training.* 2004;39(3):263.

276. Dauty M, Potiron-Josse M, Rochcongar P. Identification of previous hamstring muscle injury by isokinetic concentric and eccentric torque measurement in elite soccer player. *Isokinetics and Exercise Science.* 2003;11(3):139-144.

277. Orchard J, Seward H. Epidemiology of injuries in the Australian Football League, seasons 1997–2000. *British Journal of Sports Medicine.* 2002;36(1):39-44.

278. Brooks JH, Fuller C, Kemp S, Reddin DB. Epidemiology of injuries in English professional rugby union: part 1 match injuries. *British Journal of Sports Medicine.* 2005;39(10):757-766.

279. Brooks JH, Fuller C, Kemp S, Reddin DB. Epidemiology of injuries in English professional rugby union: part 2 training Injuries. *British Journal of Sports Medicine.* 2005;39(10):767-775.

280. Meeuwisse WH, Sellmer R, Hagel BE. Rates and risks of injury during intercollegiate basketball. *The American Journal of Sports Medicine.* 2003;31(3):379-385.

281. Malliaropoulos N, Papacostas E, Kiritsi O, et al. Posterior thigh muscle injuries in elite track and field athletes. *The American Journal of Sports Medicine.* 2010;38(9):1813-1819.

282. Brukner P, Nealon A, Morgan C, Burgess D, Dunn A. Recurrent hamstring muscle injury: applying the limited evidence in the professional football setting with a seven-point programme. *British Journal of Sports Medicine.* 2014;48:929-938.

283. Liu H, Garrett WE, Moorman CT, Yu B. Injury rate, mechanism, and risk factors of hamstring strain injuries in sports: A review of the literature. *Journal of Sport and Health Science.* 2012;1(2):92-101.

284. Orchard J, Best TM. The management of muscle strain injuries: an early return versus the risk of recurrence. *Clinical Journal of Sport Medicine.* 2002;12(1):3-5.

285. Sherry MA, Best TM. A comparison of 2 rehabilitation programs in the treatment of acute hamstring strains. *Journal of Orthopaedic & Sports Physical Therapy.* 2004;34(3):116-125.

286. Ueblacker P, Mueller-Wohlfahrt HW, Ekstrand J. Epidemiological and clinical outcome comparison of indirect (‘strain’) versus direct (‘contusion’) anterior and posterior thigh muscle injuries in male elite football players: UEFA Elite League study of 2287 thigh injuries (2001–2013). *British Journal of Sports Medicine.* 2015;49(22):1461-1465.

287. Cejudo A, Sainz de Baranda P, Ayala F, Santonja F. Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players. *Physical Therapy in Sport.* 2015;16(2):107-113.

288. Verrall GM, Kalairajah Y, Slavotinek J, Spriggins A. Assessment of player performance following return to sport after hamstring muscle strain injury. *Journal of Science and Medicine in Sport.* 2006;9(1):87-90.
289. Yu B, Queen RM, Abbey AN, Liu Y, Moorman CT, Garrett WE. Hamstring muscle kinematics and activation during overground sprinting. *Journal of Biomechanics.* 2008;41(15):3121-3126.
290. Forbes H, Bullers A, Lovell A, McNaughton LR, Polman R, Siegler JC. Relative torque profiles of elite male youth footballers: effects of age and pubertal development. *International Journal of Sports Medicine.* 2009;30(08):592-597.
291. Zvijac JE, Toriscelli TA, Merrick S, Kiebzak GM. Isokinetic concentric quadriceps and hamstring strength variables from the NFL Scouting Combine are not predictive of hamstring injury in first-year professional football players. *The American Journal of Sports Medicine.* 2013;41(7):1511-1518.
292. Ayala F, De Ste Croix M, Sainz de Baranda P, Santonja F. Absolute reliability of hamstring to quadriceps strength imbalance ratios calculated using peak torque, joint angle-specific torque and joint ROM-specific torque values. *International Journal of Sports Medicine.* 2012;33(11):909-916.
293. Ekbal A. Improvement of prediction accuracy using discretization and voting classifier. Paper presented at: Pattern Recognition, 2006. ICPR 2006. 18th International Conference on2006.
294. Bouckaert RR. Bayesian network classifiers in weka for version 3-5-7. *Artificial Intelligence Tools.* 2008;11(3):369-387.
295. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research.* 2002;16:321-357.
296. Chawla NV, Lazarevic A, Hall LO, Bowyer KW. SMOTEBoost: Improving prediction of the minority class in boosting. Paper presented at: *European Conference on Principles of Data Mining and Knowledge Discovery*. 2003.
297. Seiffert C, Khoshgoftaar TM, Van Hulse J, Napolitano A. RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans.* 2010;40(1):185-197.
298. Wang S, Yao X. Diversity analysis on imbalanced data sets by using ensemble models. Paper presented at: *Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium* on 2009.
299. Gibson EJ, Pick AD. *An ecological approach to perceptual learning and development.* Oxford University Press, USA; 2000.
300. Ramírez-Campillo R, Burgos CH, Henríquez-Olguín C, et al. Effect of unilateral, bilateral, and combined plyometric training on explosive and endurance performance of young soccer players. *The Journal of Strength & Conditioning Research.* 2015;29(5):1317-1328.
301. Gribble PA, Hertel J, Plisky P. 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.* 2012;47(3):339-357.
302. Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the star excursion balance tests in detecting reach deficits in subjects with chronic ankle instability. *Journal of Athletic Training.* 2002;37(4):501.

303. Earl JE, Hertel J, Denegar CR. Patterns of dynamic malalignment, muscle activation, joint motion, and patellofemoral-pain syndrome. *Journal of Sport Rehabilitation.* 2005;14(3):216-233.

304. Herrington L, Hatcher J, Hatcher A, McNicholas M. A comparison of Star Excursion Balance Test reach distances between ACL deficient patients and asymptomatic controls. *The Knee.* 2009;16(2):149-152.

305. Clagg S, Paterno MV, Hewett TE, Schmitt LC. Performance on the modified star excursion balance test at the time of return to sport following anterior cruciate ligament reconstruction. *Journal of Orthopaedic & Sports Physical Therapy.* 2015;45(6):444-452.

306. Hertel J, Miller SJ, Denegar CR. Intratester and intertester reliability during the Star Excursion Balance Tests. *Journal of Sport Rehabilitation.* 2000;9(2):104-116.

307. Butler RJ, Southers C, Gorman PP, Kiesel KB, Plisky PJ. Differences in soccer players' dynamic balance across levels of competition. *Journal of Athletic Training.* 2012;47(6):616-620.

308. Paillard T, Noe F, Riviere T, Vincent M. Postural performance and strategy in the unipedal stance of soccer players at different levels of competition. *Journal of Athletic Training.* 2006;41(2):172.

309. Bressel E, Yonker JC, Kras J, Heath EM. Comparison of static and dynamic balance in female collegiate soccer, basketball, and gymnastics athletes. *Journal of Athletic Training.* 2007;42(1):42.

310. Pfirrmann D, Herbst M, Ingelfinger P, Simon P, Tug S. Analysis of injury incidences in male professional adult and elite youth soccer players: a systematic review. *Journal of Athletic Training.* 2016;51(5):410-424.

311. Filipa A, Byrnes R, Paterno MV, Myer GD, Hewett TE. Neuromuscular training improves performance on the star excursion balance test in young female athletes. *Journal of Orthopaedic & Sports Physical Therapy.* 2010;40(9):551-558.

312. Hiller CE, Kilbreath SL, Refshauge KM. Chronic ankle instability: evolution of the model. *Journal of Athletic Training.* 2011;46(2):133-141.

313. Hertel J, Braham RA, Hale SA, Olmsted-Kramer LC. Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability. *Journal of Orthopaedic & Sports Physical Therapy.* 2006;36(3):131-137.

314. Kwon YJ, Park SJ, Jefferson J, Kim K. The effect of open and closed kinetic chain exercises on dynamic balance ability of normal healthy adults. *Journal of Physical Therapy Science.* 2013;25(6):671-674.

315. Leetun DT, Ireland ML, Willson JD, Ballantyne BT, Davis IM. Core stability measures as risk factors for lower extremity injury in athletes. *Medicine & Science in Sports & Exercise.* 2004;36(6):926-934.

316. Zazulak BT, Hewett TE, Reeves NP, Goldberg B, Cholewicki J. Deficits in neuromuscular control of the trunk predict knee injury risk: a prospective biomechanical-epidemiologic study. *The American Journal of Sports Medicine.* Jul 2007;35(7):1123-1130.

317. Malloy P, Morgan A, Meinerz C, Geiser CF, Kipp K. Hip external rotator strength is associated with better dynamic control of the lower extremity during landing tasks. *The Journal of Strength & Conditioning Research.* 2016;30(1):282.

318. Sugimoto D, Mattacola CG, Mullineaux DR, Palmer TG, Hewett TE. Comparison of isokinetic hip abduction and adduction peak torques and ratio between sexes. *Clinical Journal of Sport Medicine: official journal of the Canadian Academy of Sport Medicine.* 2014;24(5):422.
319. Smith HC, Vacek P, Johnson RJ, et al. Risk factors for anterior cruciate ligament injury: a review of the literature—part 2: hormonal, genetic, cognitive function, previous injury, and extrinsic risk factors. *Sports Health.* 2012;4(2):155-161.
320. Dick R, Putukian M, Agel J, Evans TA, Marshall SW. Descriptive epidemiology of collegiate women's soccer injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2002–2003. *Journal of Athletic Training.* 2007;42(2):278.
321. O'Kane JW, Tencer A, Neradilek M, Polissar N, Sabado L, Schiff MA. Is knee separation during a drop jump associated with lower extremity injury in adolescent female soccer players? *The American Journal of Sports Medicine.* 2016;44(2):318-323.
322. Mackenzie T, Herrington L, Funk L, Horlsey I, Cools A. Sport Specific Adaptation in Rotation Range of Motion in the Elite Golfer's Shoulder. *Journal of Athletic Enhancement*. 2016;5:2.
323. Ellenbecker TS, Roetert EP, Bailie DS, Davies GJ, Brown SW. Glenohumeral joint total rotation range of motion in elite tennis players and baseball pitchers. *Medicine & Science in Sports & Exercise.* 2002;34(12):2052-2056.
324. Cejudo A, Sainz de Baranda P, Ayala F, Santonja F. Perfil de flexibilidad de la extremidad inferior en jugadores senior de balonmano. *Cuadernos de Psicología del Deporte.* 2014;14(2):111-120.
325. Kleinberg S. *Why: A guide to finding and using causes.* " O'Reilly Media, Inc."; 2015.