©Journal of Sports Science and Medicine (2026) **25**, 172-194
http://www.jssm.org DOI: https://doi.org/10.52082/jssm.2026.172

Review article

Machine Learning Applications in Non-Contact Lower Limb Sports Injury Prediction: A Systematic Review

Jin Yuan [1], Quanwen Zeng [1], Anjie Wang [1], Yong Zhang [1]✉ and Jun Li [2]✉
[1] School of Physical Education, Anhui Polytechnic University, Wuhu, 241000, Anhui, China
[2] School of Athletic Performance, Shanghai University of Sport, Shanghai, 200438, China

Abstract
Non-contact Lower limb sports injuries represent some of the most prevalent and impactful conditions within athletic populations, prompting increasing interest in predictive approaches that can inform prevention and rehabilitation strategies. With its capacity to manage high-dimensional and complex datasets, machine learning (ML) has emerged as a promising tool for injury risk prediction. This systematic review, conducted in accordance with PRISMA 2020 guidelines, synthesized evidence from studies retrieved through Web of Science, PubMed, and SPORTDiscus (EBSCO). The literature search was conducted on January 20, 2025.Following independent screening and risk of bias assessment using the PROBAST tool, 15 studies were included from an initial pool of 92. The majority of study populations comprised adult athletes, with basketball and football (soccer) being the most frequently investigated sports. Random Forest and logistic regression were the most commonly applied algorithms, while tree-based approaches yielded the strongest predictive performance in 6 studies. Across 14 studies, area under the curve (AUC) values were reported, with one CHAID-based decision tree achieving the highest performance (AUC = 0.91), and sensitivity values reaching up to 0.92 in eight studies. Importantly, model interpretability was addressed in 87% of included studies, underscoring its emerging importance for clinical translation. Overall, ML exhibits considerable potential in predicting non-contact lower-limb injuries, but its practical value depends on achieving a balance between accuracy, transparency, and reliability. Future research should emphasize the integration of multi-source data and large-scale prospective validation to advance the translation of ML models into precision injury prevention and rehabilitation practice.

Key words: Predictive analytics, sports medicine, risk factors, risk assessment, rehabilitation, predictive models.

Introduction

Non-contact lower limb injuries constitute a notable subset of musculoskeletal sports injuries and are of particular importance because they typically arise in the absence of external impact, making them more challenging to predict and more closely linked to modifiable intrinsic and biomechanical risk factors (Belkhelladi et al., 2025; Whittaker et al., 2025). Across youth and adult athletes, such injuries frequently lead to time loss and more than half of anterior cruciate ligament (ACL) injuries in team sports arise from non-contact mechanisms such as cutting or sudden deceleration (Chia et al., 2022; Guan et al., 2021). These injuries are particularly prevalent in sports with repeated high-intensity directional changes—most notably soccer, basketball and rugby—where epidemiological studies consistently report elevated non-contact injury rates (Achenbach et al., 2021; Ekstrand et al., 2011; Evans et al., 2024; López-Valenciano et al., 2020). In elite soccer, for example, over 90% of lower-limb muscle injuries occur through non-contact mechanisms (Ekstrand et al., 2011). Importantly, non-contact injuries are broadly considered preventable, with evidence showing that neuromuscular and strength-focused injury-prevention programs can substantially reduce their incidence (Al Attar et al., 2017; Rössler et al., 2018; Webster and Hewett, 2018; Yu and Garrett, 2007). Beyond their high incidence, non-contact injuries also impose meaningful economic burdens; in the Australian Football League, the annual financial loss per club reaches AUD$188k to 333k, with missed matches due to hamstring strain injuries, predominantly non-contact, increasing by 71% between 2003 and 2012 (Hickey et al., 2014; Lu et al., 2021).

The lack of consensus on the risk factors for non-contact lower limb sports injuries poses a considerable challenge to accurately identifying their underlying causes. Traditional univariate analytical approaches are inherently limited, as their conclusions are often fragmented and fail to account for the complex interactions among multidimensional factors within dynamic sporting environments(Ruddy et al., 2019). Increasing evidence indicates that injuries emerge from nonlinear interactions among physiological, biomechanical, psychological, and environmental variables rather than from any single determinant(Green et al., 2020; Liveris, 2025). This recognition has prompted a shift from linear, single-cause analyses toward more comprehensive and systematic modeling approaches (Bittencourt et al., 2016), enabling identification of critical combinations of risk factors and providing a stronger scientific foundation for individualized injury prediction and prevention strategies.

In recent years, the field of sports science has increasingly adopted machine learning (ML) approaches to uncover latent patterns within large-scale and complex datasets, demonstrating substantial utility in areas such as competition outcome prediction, performance optimization, and tactical decision-making(Horvat and Job, 2020; Hubáček et al., 2019; Ou-Yang et al., 2025; Sampaio et al., 2024; Watson et al., 2021). These advances are gradually reshaping the landscape of sports medicine. However, conventional statistical techniques (primarily logistic regression) struggle to model nonlinear relationships and are

Received: 12 September 2025 / Accepted: 08 December 2025 / Published (online): 01 March 2026

prone to biased performance when faced with the pronounced class imbalance typical of prospective injury datasets. As a result, these models frequently classify the majority of non-injury cases correctly while showing substantially reduced sensitivity and limited discriminative capacity for the minority injury outcomes (Lopez-Valenciano et al., 2018; Oliver et al., 2020; Rossi et al., 2018; Ruddy et al., 2018; Ruiz-Perez et al., 2021). By comparison, ML techniques can accommodate nonlinear relationships and complex feature interactions within high-dimensional, multimodal datasets, enabling a more nuanced characterization of injury-related patterns. While not uniformly superior across all applications, ML approaches have shown potential to yield improved sensitivity and more informative risk stratification in certain contexts (Ayala et al., 2019). Furthermore, ML offers a unique advantage in its ability to integrate a broad range of athlete-specific variables, including sport experience, training load characteristics, biological sex, performance level, prior injury history, and sport-specific biomechanical demands, into unified predictive frameworks (Bogaert et al., 2022; Musat et al., 2024; Rommers et al., 2020). This capacity to model complex, individualized risk profiles is especially relevant for non-contact lower limb injuries, which arise from multifactorial and predominantly intrinsic mechanisms. Although challenges remain due to substantial inter-individual variation in tissue tolerance and adaptive capacity (Nassis et al., 2023), continued progress in multimodal data fusion, feature engineering, and rigorous model validation is steadily enhancing the precision and practical relevance of ML-based injury risk estimation. These developments are expected to support more dependable individualized assessments and contribute to more targeted, evidence-informed prevention strategies (Bartlett et al., 2017; Rossi et al., 2018; Wilkerson et al., 2018; Willy, 2018).

Recent reviews have explored machine learning applications in sports injury prediction, including the systematic review by Van Eetvelde et al. (2021), the scoping review by Leckey et al. (2025), and the narrative review by Yuan et al. (2025). These studies provided important overviews of general ML developments and highlighted shared challenges such as heterogeneous data sources, inconsistent injury definitions, small sample sizes, and limited interpretability and external validation, but they largely evaluated ML at a global level across multiple injury types and body regions. Leckey et al. (2025) provided a broad evidence synthesis of ML methods across sports but did not perform an anatomically or mechanism-focused analysis, and Van Eetvelde et al. (2021) emphasized the need for future work to focus on interpretable ML models and injury-specific analyses, yet their review did not provide region- or mechanism-targeted evaluations. Yuan et al. (2025) structured their narrative around the workflow of injury prediction model development, highlighting methodological challenges encountered during model development, such as data preprocessing, feature selection, and model evaluation, but without stratifying findings by anatomical region or injury mechanism. Therefore, a focused, up-to-date synthesis that examines ML applications specifically for non-contact lower-limb injuries, attending to data modalities, class-imbalance strategies, temporal prediction windows, injury-type heterogeneity, and interpretability practices, is warranted to generate more actionable, domain-relevant guidance. To clarify how machine learning methods interface with injury mechanisms, data modalities, and prediction tasks, we present a conceptual framework summarizing the key components of ML-based non-contact lower-limb injury prediction (Figure 1). This framework also serves to situate the scope of the present review within the broader methodological landscape.

In light of these developments and the growing need for precise, individualized risk assessment, the present systematic review aims to synthesize current evidence on ML applications for lower limb non-contact injury prediction and address the following objectives:

(1) Summarize the main ML approaches employed in injury prediction and their methodological characteristics;

(2) Evaluate their effectiveness in terms of predictive performance; and

(3) Examine the role of interpretability techniques in existing studies and assess their implications for clinical translation and practical application.

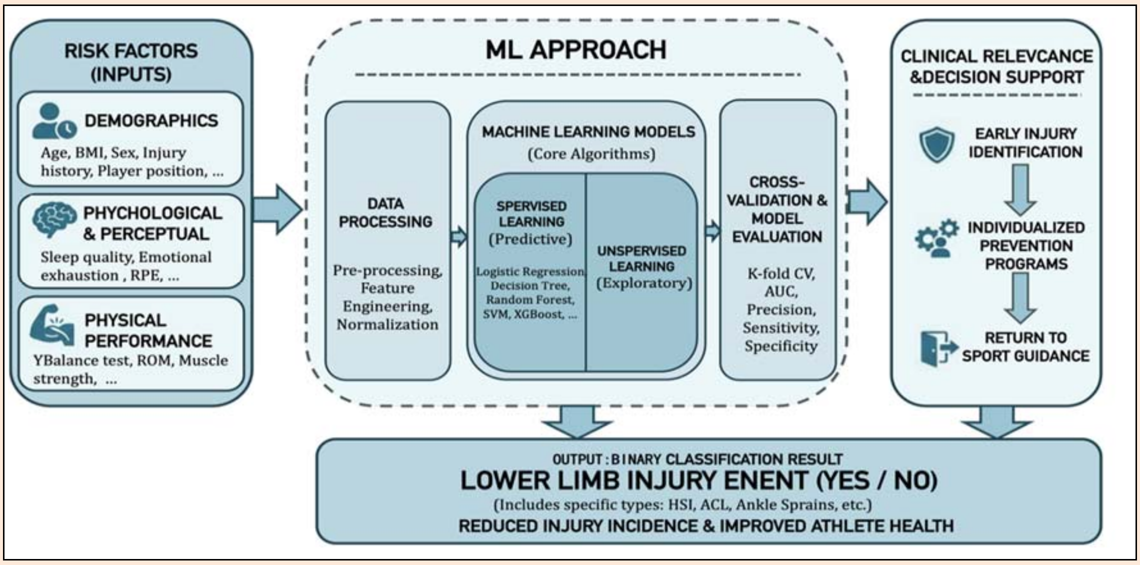

Figure 1. Conceptual framework of machine learning approaches for lower limb sports injury prediction.

Methods

Study design

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines(Page et al., 2021). The review protocol was prospectively registered in PROSPERO (ID: CRD420251070408).

Search strategy

As of January 20, 2025, a comprehensive literature search was performed across three electronic bibliographic databases: Web of Science, PubMed, and SPORTDiscus (EBSCO), using the following search terms: ('athletic injuries' OR 'sports injuries') AND ('machine learning' OR 'transfer learning') AND ('lower extremity' OR 'lower limbs') (the complete search strategy is available in the Supplementary Table 1). Additionally, three reviewers (JY, YZ, and QZ) independently conducted the database search and cross-checked the reference lists of relevant studies.

Inclusion and exclusion criteria

Studies were included if they met the following criteria: (1) published in English in peer-reviewed journals; (2) original research articles; (3) applied ML techniques to predict non-contact lower limb injuries in humans; and (4) involved athletes or physically active populations, with study designs encompassing prospective, retrospective, or cross-sectional approaches; and (5) published within the last ten years (January 2015 to January 20, 2025). Exclusion criteria were as follows: (1) full text not available in English; (2) studies that did not employ ML methods for non-contact lower limb injury prediction (e.g., those limited to traditional regression analyses); (3) studies that did not report the number of injury cases; and (4) review articles, conference abstracts, or editorials.

Study selection and data extraction

All records were screened following a predefined protocol. Grey literature (e.g., theses, dissertations, non-peer-reviewed reports) was excluded a priori, as the review focused exclusively on peer-reviewed scientific evidence. Two independent reviewers (JY and QZ) screened the titles and abstracts of all retrieved studies to determine eligibility based on the predefined inclusion criteria. Inter-rater agreement during screening was assessed using Cohen's κ coefficient, with discrepancies resolved through discussion with a third reviewer (YZ) until consensus was reached. For the studies that met the inclusion criteria, the reviewers independently extracted relevant data using a predesigned standardized data extraction form(Fernandez-Felix et al., 2023), followed by cross-checking to ensure accuracy. The extracted information included: (1) study characteristics (study design, authors, year of publication, study population, and sample size); (2) machine learning methodology (model type, feature variables, data preprocessing methods, training strategies, performance metrics, and interpretability); and (3) injury-related information (type of injury, anatomical location, number of injury events, and injury definition). If certain information was not reported in a study, it was recorded as "not reported."

Risk of bias and applicability assessment

Two independent reviewers (JY and QZ) assessed the risk of bias for the included studies, with discrepancies resolved through arbitration by a third reviewer (YZ) until consensus was reached. Assessment was conducted using the Prediction model Risk Of Bias Assessment Tool (PROBAST)(Wolff et al., 2019), which evaluates four domains: participants, predictors, outcomes, and analysis, comprising 20 signaling questions to determine both domain-specific and overall risk of bias (low, unclear, high). Key considerations included participant representativeness and inclusion criteria; predefinition and reliability of predictors; objectivity and blinding of outcome assessment; and analysis-related issues such as model overfitting, data leakage, failure to address class imbalance, and validation strategy. Applicability was assessed based on the relevance of study participants, predictors, and outcomes to the review question.

Results

The outcomes of the search strategy and study selection process are illustrated in Figure 2. A total of 86 potentially relevant studies were initially identified through systematic searches of the Web of Science, PubMed, and SPORTDiscus databases. An additional 6 articles were retrieved through manual searching, yielding a total of 92 records. After removing 15 duplicates, 77 unique articles remained. Following title and abstract screening, 42 studies were deemed eligible based on the inclusion criteria. Inter-rater reliability for screening was substantial (Cohen's $\kappa = 0.72$). The results from the independent screenings were subsequently consolidated, and any discrepancies were resolved through discussion among the three reviewers (JY, YZ, and QZ). Ultimately, 15 studies were included in the final review. The descriptive characteristics of the included studies are summarized in Supplementary Table 2; Supplementary Table 3 provides the ML and statistical definitions referenced therein.

Risk of bias and applicability assessment

Among the 15 included studies, four were rated as having a low overall risk of bias, eight as unclear, and three as high risk (see Figure 3 and Figure 4). Bias was predominantly observed in the analysis domain, with key issues including insufficient sample sizes, unreported sensitivity and specificity, and inadequate handling and reporting of missing data. Regarding applicability, nine studies were rated as low concern, four as unclear, and two as high concern. The main applicability limitations were unclear inclusion and exclusion criteria and limited relevance of the predicted outcomes to actual lower-limb muscle injury risk, which may reduce the practical utility of the findings for real-world sports or clinical prevention.

Sporting contexts and participant characteristics

The distribution of publications by year showed a trend of initial growth followed by a subsequent decline: 2018 (n = 2), 2019 (n = 2), 2020 (n = 2), 2021 (n = 2), 2022 (n = 4), and 2023 (n = 3). Regarding sport type, four studies focused on soccer(Ayala et al., 2019; Javier Robles-Palazon et al., 2023; Kolodziej et al., 2023; Oliver et al., 2020), three on basketball(Huang et al., 2022; Huang et al., 2023; Lu et al., 2022), and one study each on football(Ruddy et al., 2018), futsal(Ruiz-Perez et al., 2021), and military personnel(Connaboy et al., 2019). Five studies included multiple sport populations(Bogaert et al., 2022; Henriquez et al., 2020; Jauhiainen et al., 2022; Jauhiainen et al., 2021; Lopez-Valenciano et al., 2018).

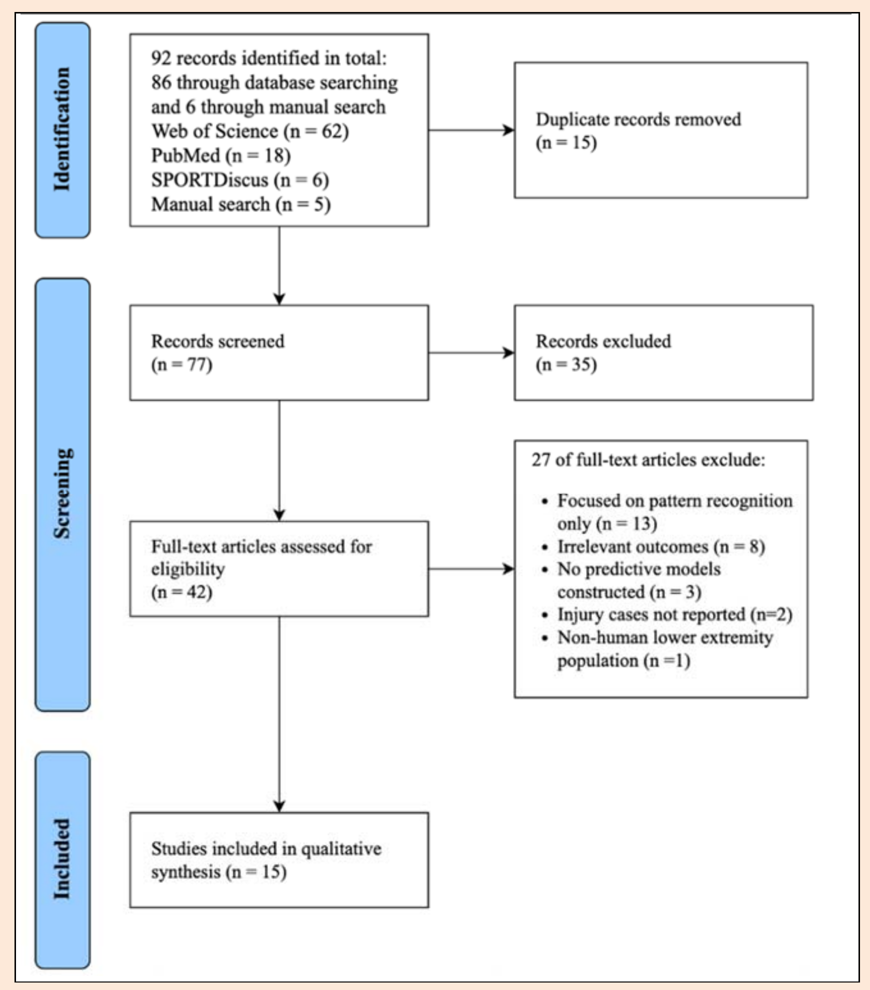

Figure 2. **PRISMA flow diagram for the description of the overall process.**

Author, Year	Risk of Bias				Applicability			Overall	
	1. Participants	2. Predictors	3. Outcome	4. Analysis	1. Participants	2. Predictors	3. Outcome	Risk of Bias	Applicability
(1) Lopez-Valenciano, 2018	+	+	+	+	+	+	+	+	+
(2) Ruddy, 2018	+	+	+	?	+	+	+	?	+
(3) Ayala, 2019	+	+	+	?	+	+	+	?	+
(4) Connaboy , 2019	?	+	+	?	?	+	+	?	?
(5) Henriquez, 2020	-	+	?	?	-	+	+	-	-
(6) Oliver, 2020	+	+	+	?	+	+	+	?	+
(7) Jauhiainen, 2021	+	+	+	?	+	+	+	?	+
(8) Ruiz-Perez, 2021	+	+	+	+	+	+	+	+	+
(9) Bogaert, 2022	+	+	+	?	+	+	+	?	+
(10) Huang, 2022	+	+	+	-	?	+	+	-	?
(11) Lu, 2022	?	+	?	?	+	+	?	?	?
(12) Huang, 2023	?	+	+	-	-	+	+	-	-
(13) Javier Robles-Palazon, 2023	+	+	+	+	+	+	+	+	+
(14) Kolodziej, 2023	+	+	+	+	+	+	+	+	+
(15) Jauhiainen, 2022	+	+	+	?	?	+	+	?	?

Figure 3. **Risk of bias and applicability assessment.**

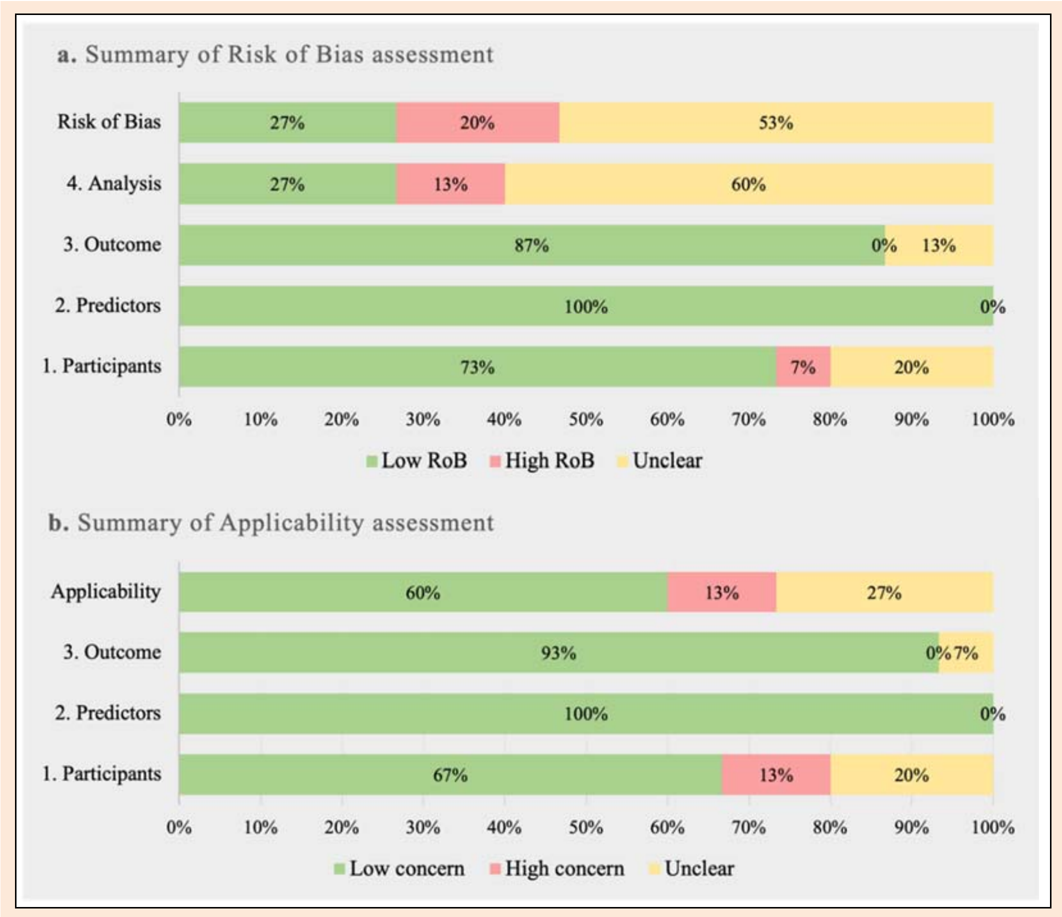

Figure 4. Summary of risk of bias and applicability assessment within the study.

Table 1. Gender and type of participants.

Author(year)	Participant	Type
Lopez-Valenciano et al (2018)	Male (132)	Adult Athletes
Ruddy et al (2018)	Male (362)	Adult Athletes
Ayala et al (2019)	Male (96)	Adult Athletes
Connaboy et al (2019)	Not reported	Military personnel
Henriquez et al (2020)	Mix (122, Male: 71, Female: 51)	Adult Athletes
Oliver et al (2020)	Not reported	Adolescent Athletes
Jauhiainen et al (2021)	Mix (314, Male:152, Female: 162)	Adolescent Athletes
Ruiz-Perez et al (2021)	Mix (139, Male: 72, Female: 67)	Adult Athletes
Bogaert et al (2022)	Mix (204, Male: 141, Female: 63)	Recreationally active individuals
Jauhiainen et al (2022)	Female (791)	Adult Athletes
Huang et al (2022)	Female (16)	Adolescent Athletes
Lu et al (2022)	Male (2103)	Adult Athletes
Huang et al (2023)	Female (17)	Adolescent Athletes
Javier Robles-Palazon et al (2023)	Male (260)	Adolescent Athletes
Kolodziej et al (2023)	Male (56)	Adolescent Athletes

Sample sizes ranged from 16 to 2103 participants (Table 1). In terms of sex distribution, six studies recruited only male participants (40%), four included mixed-gender participants (27%), three recruited only female participants (20%), and two did not report participant sex. Regarding age or population characteristics, seven studies involved adult athletes (47%), six involved adolescent athletes, one included recreationally active individuals, and one involved military personnel.

Data characteristics analysis

Among the 15 included studies, the majority (n = 11, 73%) evaluated the predictive ability of machine learning models for injuries occurring in any region of the lower limb, while the remaining studies focused on specific anatomical sites, including hamstring strain injuries (HSI, n = 2) (Ayala et al., 2019; Ruddy et al., 2018) , knee (n = 1)(Jauhiainen et al., 2022)and one study predicting both knee and ankle injuries (Jauhiainen et al., 2021). Most studies (n = 10) targeted traumatic injuries, three addressed overuse injuries (Bogaert et al., 2022; Huang et al., 2022; Huang et al., 2023) , and two did not specify the injury mechanism (Connaboy et al., 2019; Henriquez et al., 2020) (Table 2). Regarding predictor variables, the most frequently collected features could be classified into three major domains: demographics and injury history, psychological and perceptual variables, and physical performance measures. Thirteen studies (87%) incorporated demographic information such as age, height, weight, competitive level, and prior injury history, with nine studies explicitly including previous injury as a predictor. Five studies assessed psychological and perceptual factors, most commonly sleep quality (n = 4)(Ayala et al., 2019; Huang et al., 2022; Lopez-Valenciano et al., 2018; Ruiz-Perez et al., 2021), alongside other constructs such as sport anxiety, team cohesion, and stress levels. Less commonly, innovative predictors such as urinary biomarkers (n = 2)(Huang et al., 2022; Huang et al., 2023) and match performance indicators (n = 1)(Lu et al., 2022) were also reported.

Table 2. Data analysis characteristics.

Author (Year)	Lower Extremity Location	Predictor Variables	Data Pre-processing	Feature Selection/ Dimensionality Reduction	Training Strategy
Lopez-Valenciano et al. (2018)	Any muscle (traumatic)	- Demographics and Injury History (Age, BMI, Injury history, Sleep, Level of play) - Psychological and Perceptual Variables (Sport devaluation, Sleep quality) - Physical Performance Measures (YBalance test, Core control, ROM of hip, Isometric strength)	-Data imputation -Weka software	NR	- SMOTE -5-fold cross-validation
Ruddy et al. (2018)	HSI (traumatic)	- Demographics and Injury History (Age, Height, Mass. Playing Position, History HIS, History ACL) - Physical Performance Measures (Peak Hamstring Force Right, Peak Hamstring Force Left, Hamstring Force Imbalance)	Data normalized (Z-score)	NR	- SMOTE - 10-fold cross-validation
Ayala et al. (2019)	HSI (traumatic)	- Demographics and Injury History (Age, History of HSI last season, Maximal level of play achieved) - Psychological and Perceptual Variables (Sleep quality, Physical/emotional exhaustion, Reduced sense of accomplishment) - Physical Performance Measures (Dynamic postural control, Isometric hip abduction and adduction strength, Lower extremity joint ROMs)	Data imputation	NR	- SMOTE - 3-fold cross-validation
Connaboy et al. (2019)	Any region (NR)	- Demographics (Age, Boday fat, Weight) - Physical Performance Measures (Peak anaerobic power, Mean anaerobic power, Knee active extension)	Not Reported	NR	Leave-one out cross validation
Henriquez et al. (2020)	Any region (NR)	- Demographics (Weight, Height, Gender, Age) - Physical Performance Measures (Eyes Open Balance Test Composite Score, DPSI Composite Score, Straight Leg Raise, Active Knee Extension, Ankle Dorsiflexion Strength)	Data normalized (Z-score)	Mean Decrease Accuracy	5-fold cross-validation
Oliver et al. (2020)	Any region (traumatic)	-Demographic (Age, BMI, Height); -Physical Performance Measures (Maturity-Offset, 75%Hop L PVGRF, 75%Hop R PVGRF)	Weka software	NR	- Cost-sensitive learning - 5-fold cross-validation
Jauhiainen et al. (2021)	Knee, Ankle (traumatic)	- Demographic and Injury History (Age, Sex, BMI, Previous ACL, Family ACL history) - Physical Performance Measures (KT1000 (dominant leg), hip flexion peak (dominant leg), medial knee displacement (both legs), vertical ground reaction force (vGRF) (both legs))	-Data imputation -Data normalized (Z-score)	Expert-based feature selection	10-fold cross-validation
Ruiz-Perez et al. (2021)	Any region (traumatic)	- Demographics and Injury History (Player position, Current level of play, Dominant leg, Sex, Age) - Psychological and Perceptual Variables (Physical/emotional exhaustion, Sleep quality) - Physical Performance Measures (PosteroLateral, Y-Balance-Composite)	-Data imputation -Weka software	Attribute Selected Classifier	-Under-sampling Bagging -5-fold cross-validation
Bogaert et al. (2022)	Any region (overuse)	- Demographics and Injury History (Gender, Weight, Height, Previous injuries) - Physical Performance Measures (Root-mean-square ratio, Step regularity, Stride regularity, Sample entropy)	Data normalized (Min-Max Scaling)	PCA (Principal Component Analysis)	-Cost-sensitive learning -Internal Cross-Validation

NR, not reported; HSI, hamstring strain injuries.

Table 2. Continue…

Author (Year)	Lower Extremity Location	Predictor Variables	Data Pre-processing	Feature Selection/ Dimensionality Reduction	Training Strategy
Jauhiainen et al. (2022)	Knee (traumatic)	- Demographics and Injury History (Age, Body mass, Previous ACL) - Physical Performance Measures (Single leg drop jump knee, Jump hip flex max)	-Data imputation -Data normalized	NR	- SMOTE -5-fold cross-validation
Huang et al. (2022)	Any region (overuse)	- Psychological and Perceptual Variables (Sleep Quality, Muscle Soreness, Stress Levels) - Physical Performance Measures (Squat 1RM, 15 m × 17 Shuttle Run, 5.8 m × 6 Shuttle Run) - Physiological Status Indicators (Urine Protein, Urobilinogen, Urine pH, Urine Specific Gravity)	-Data imputation -Data normalized (Z-score)	NR	- SMOTE - 10-fold cross-validation
Lu et al. (2022)	Any Muscle (traumatic)	- Demographics and Injury History (Recent hamstring injury, Recent back injury, Age) - Game Performance Metrics (Field goal percentage, 3-point shots made per game, 3-point shots attempted per game, Usage percentage, Offensive win share, Defensive win share)	Not Reported	RFE (Recursive Feature Elimination)	10-fold cross-validation
Huang et al. (2023)	Any region (overuse)	- Psychological and Perceptual Variables (RPE: Ratings of Perceived Exertion) - Physical Performance Measures (Double under, Squat, Bench press, Shuttle run, Sprint) - Physiological Status Indicators (Instantaneous Heart Rate, Heart Rate Recovery, Protein, pH)	-Data imputation -Data normalized (Z-score)	LDA (Linear Discriminant Analysis)	- SMOTE - 5-fold cross-validation
Javier Robles-Palazon et al. (2023)	Soft tissue (traumatic)	- Demographics and Injury History (Age, BMI, Injury history) - Psychological and Perceptual Variables (Anxiety, Motivation, Team cohesion) - Physical Performance Measures (Joint Range of Motion (ROM), Balance and Stability, Functional Performance Measures)	-Data imputation -Weka software	Attribute Selected Classifier	- Under-sampling Bagging - 5-fold cross-validation
Kolodziej et al. (2023)	Any region (traumatic)	- Demographics (Age, Height, Weight) - Physical Performance Measures (Postural Control and Balance, Strength Measures, Joint Kinematics, Joint Moments) - Ground Reaction Forces (Peak vGRF)	Data normalized	LASSO (Least Absolute Shrinkage and Selection Operator)	15-fold cross-validation

NR, not reported; HSI, hamstring strain injuries.

In terms of data preprocessing, 13 studies reported explicit procedures. The most common steps included data normalization (n = 8, using either Z-score standardization or Min-Max scaling) to harmonize feature scales, and data imputation to handle missing values. Additionally, four studies used the Weka software package for preprocessing (Javier Robles-Palazon et al., 2023; Lopez-Valenciano et al., 2018; Oliver et al., 2020; Ruiz-Perez et al., 2021), including imputation and discretization. Two studies did not report any preprocessing.

Feature selection or dimensionality reduction techniques were reported in eight studies (53%). Expert-based feature selection was used in one study(Jauhiainen et al., 2021), while the Attribute Selected Classifier from Weka was applied in two(Javier Robles-Palazon et al., 2023; Ruiz-Perez et al., 2021). Feature importance ranking (Mean Decrease Accuracy) was adopted in one study(Henriquez et al., 2020). Dimensionality reduction techniques included principal component analysis (PCA, n = 1)(Bogaert et al., 2022), recursive feature elimination (RFE, n = 1)(Lu et al., 2022), linear discriminant analysis (LDA, n = 1)(Huang et al., 2023), and least absolute shrinkage and selection operator (LASSO, n = 1)(Kolodziej et al., 2023). The remaining seven studies did not report any such methods.

A key characteristic of the included datasets was class imbalance. Based on the sample counts reported in the original studies, we recalculated the imbalance ratio (IR = minority/majority, where minority refers to injury cases and majority to non-injury cases). The average IR across studies was 0.35, with values ranging from 0.08 to 0.77. To address this issue, ten studies adopted imbalance-handling strategies, which could be broadly grouped into two categories: (i) resampling methods, including SMOTE (n = 4)(Ayala et al., 2019; Huang et al., 2022; Huang et al., 2023; Jauhiainen et al., 2022; Lopez-Valenciano et al., 2018; Ruddy et al., 2018) and under-sampling bagging (n = 2)(Javier Robles-Palazon et al., 2023; Ruiz-Perez et al., 2021); and (ii) cost-sensitive learning (n = 2)(Bogaert et al., 2022; Oliver et al., 2020). All included studies adopted cross-validation methods such as 5-fold, 10-fold, or leave-one-out.

Commonly used machine learning models

Among the 15 included studies, 4 (27%) employed a single ML model for predictive modeling, whereas the remaining studies compared multiple models to identify the one with optimal predictive performance (Table 3). Specifically, 1 study evaluated 2 models, 3 studies evaluated 3 models, 5 studies evaluated 4 models, and 2 studies assessed more than 4 models. Across all studies, random forest (RF) and logistic regression were the most frequently applied algorithms, each appearing in 8 studies (53%), followed by support vector machine (SVM), which were used in 7 studies. In addition, decision trees and their variants (e.g., C4.5, SimpleCart, ADTree, CHAID) were applied in 6 studies. By contrast, extreme gradient boosting (XGBoost) was less commonly used, reported in only 2 studies. Overall, tree-based models and their ensemble methods emerged as the most prevalent approaches for sports injury prediction.

Table 3. Study results characteristics.

Author (Year)	ML Algorithm Used	Best-Performing Algorithm	Model Performance	Model Interpretability (Important Injury Predictors)
Lopez-Valenciano et al. (2018)	C4.5, SimpleCart, ADTree, RF	ADTree	AUC = 0.75, Sensitivity = 0.66, Specificity = 0.69	Interpretable by Design (sport devaluation, history of muscle injury in last season)
Ruddy et al. (2018)	Naïve Bayes, LR, RF, SVM, NN	Naïve Bayes	AUC = 0.60	Not Reported
Ayala et al. (2019)	J48, SimpleCart, ADTree	ADTree	AUC = 0.84, Sensitivity = 0.78, Specificity = 0.84	Interpretable by Design (sleep quality, history of HSI last season, range of motion – passive hip flexion with knee extended)
Connaboy et al. (2019)	CHAID	DT	AUC = 0.91	Interpretable by Design (knee flexion angle asymmetry, body mass)
Henriquez et al. (2020)	RF	RF	AUC = 0.69	Interpretable by Design (hip external rotation strength, hip adductor strength, straight leg raises)
Oliver et al. (2020)	LR, DT	DT	AUC = 0.66, Sensitivity = 0.56, Specificity = 0.74	Interpretable by Design (single leg counter movement jump peak vertical ground reaction force asymmetry, body mass, leg length)
Jauhiainen et al. (2021)	RF, LR, SVM	LR	AUC = 0.65	Interpretable by Design (sex, body mass index, hamstring flexion non-dominant, KT1000 dominant)
Ruiz-Perez et al. (2021)	C4.5, ADTree, KNN, SVM	SVM	AUC = 0.77, Sensitivity = 0.66, Specificity = 0.62	Interpretable by Design (hip flexion ROM, ankle dorsiflexion ROM)
Bogaert et al. (2022)	LR, RF, SVM	SVM	Male (AUC = 0.62), Female (AUC = 0.65)	Logistic Regression (Male: vertical acceleration-derived features; Female: medial-lateral-acceleration-derived features)
Jauhiainen et al. (2022)	RF, LR, SVM	SVM	AUC = 0.63	Not Reported
Huang et al. (2022)	dFusionModel	dFusionModel	Precision = 0.93, Sensitivity = 0.92	SHAP (Minimal LENCI: stress, squat 1RM; Mild LENCI: sRPE, sleep, urine protein, urine blood)
Lu et al. (2022)	Elastic Net, RF, XGBoost, SVM, NN, LR	XGBoost	AUC = 0.84	SHAP (history of a back, quadriceps, hamstring, groin, or ankle injury; Concussion within the previous 8 weeks; Total count of previous injuries.)
Huang et al. (2023)	Cost-NN, LR, RF, XGBoost	Cost-NN	AUC = 0.86, Precision = 0.64, Sensitivity = 0.87	SHAP (hexagon agility test, three-quarter court sprint)
Javier Robles-Palazon et al. (2023)	C4.5, ADTree, SVM, KNN	SVM	AUC = 0.70, Sensitivity = 0.54, Specificity = 0.74	SHAP (knee maximum displacement (dominant leg) in the drop vertical jump, landing bilateral peak vertical ground reaction force (single-leg countermovement jump), BMI)
Kolodziej et al. (2023)	LASSO LR	LASSO LR	AUC = 0.63, Sensitivity = 0.35, Specificity = 0.79	Interpretable by Design (concentric knee extensor peak torque, hip transversal plane moment in the SLDL, COP sway)

ADTree, alternating decision tree; RF, random tree; LR, logistic regression; SVM, support vector machine; NN, neural network; CHAID, chi-square automatic interaction detection; KNN, k-nearest neighbor; XGBoost, extreme gradient boosting; dFusionModel, RF-based fusion of XGBoost submodels; Cost-NN, cost-sensitive neural network.

Best-performing machine learning models and evaluation metrics

Among the 15 included studies, four (27%) identified decision trees (DT) as the best-performing models(Ayala et al., 2019; Connaboy et al., 2019; Lopez-Valenciano et al., 2018; Oliver et al., 2020), followed by SVM (n = 4)(Bogaert et al., 2022; Jauhiainen et al., 2022; Javier Robles-Palazon et al., 2023; Ruiz-Perez et al., 2021) and logistic regression (LR, n = 2)(Jauhiainen et al., 2021; Kolodziej et al., 2023). Notably, one study employing the CHAID variant of DT reported the highest predictive performance across all studies (AUC = 0.91)(Connaboy et al., 2019). Overall, six studies (40%) demonstrated that tree-based algorithms, including RF, XGBoost, and DT variants were the most effective, underscoring their advantage in balancing interpretability, generalizability, and stability.

With respect to model evaluation, the area under the curve (AUC) was the most widely used metric, reported in 14 studies (93%). Among the studies reporting AUC, seven (47%) fell within the "poor" range (0.50 - 0.69)(Bogaert et al., 2022; Henriquez et al., 2020; Jauhiainen et al., 2022; Jauhiainen et al., 2021; Kolodziej et al., 2023; Oliver et al., 2020; Ruddy et al., 2018) (0.50 - 0.69), three (20%) were rated as "fair" range (0.70 - 0.79)(Javier Robles-Palazon et al., 2023; Lopez-Valenciano et al., 2018; Ruiz-Perez et al., 2021), another three as "good" range (8.80 - 0.89)(Ayala et al., 2019; Huang et al., 2023; Lu et al., 2022), and only one reached the "excellent" level (≥ 0.90)(Connaboy et al., 2019). The mean AUC across all studies was 0.73. In addition to AUC, sensitivity was the second most frequently reported metric, appearing in eight studies (53%), with values ranging from 0.35 to 0.92 and a mean of 0.63. Specificity was reported in six studies, ranging from 0.62 to 0.84 with a mean of 0.74, while precision was reported in only two studies.

Model interpretability

Among the 15 included studies, 13 (87%) reported interpretability analyses. Eight studies relied on inherently interpretable models ("white-box" algorithms(Belle and Papantonis, 2021)), primarily DT, LR, and RF. Five studies used post-hoc interpretability techniques, including SHAP (n = 4) and logistic regression applied to SVM (n = 1)(Bogaert et al., 2022).

Across the studies conducting interpretability analysis, body mass index (BMI) and previous injury history were consistently identified as important predictors across multiple studies (BMI: 4 studies(Connaboy et al., 2019; Jauhiainen et al., 2021; Javier Robles-Palazon et al., 2023; Oliver et al., 2020); previous injury history: 3 studies(Ayala et al., 2019; Lopez-Valenciano et al., 2018; Lu et al., 2022)). In addition, biomechanical features—particularly range of motion (ROM), muscle strength, and neuromuscular control—were identified as relevant predictors in several studies. These variables were primarily obtained from laboratory-based assessments using isolated screening tests, such as isokinetic or isometric strength testing(Jauhiainen et al., 2021; Kolodziej et al., 2023), goniometric or motion-capture–based ROM evaluation(Ayala et al., 2019) , and balance or perturbation tasks to assess neuromuscular control. In fewer cases, validated field-based protocols (e.g., the ROM-Sport battery(Ruiz-Perez et al., 2021)) were used to capture these capacities in applied settings.

Discussion

This systematic review synthesized 15 studies investigating ML approaches for non-contact lower limb injury prediction. Overall, tree-based algorithms were the most frequently applied and often achieved the highest predictive performance, with one study using the decision-tree variant CHAID reaching an AUC of 0.91 (Connaboy et al., 2019), exceeding the mean AUC (0.73) across studies by 25%. While AUC was the primary evaluation metric in most studies (93%), sensitivity values, reported in a subset of studies, varied widely (0.35-0.92, mean = 0.63), highlighting differences in models' ability to identify actual injury cases.

A notable feature of the included studies is that most (79%, n = 11) generalized the prediction target to "any lower limb injury event." Although this approach increases statistical power in smaller datasets, it reduces clinical specificity because different injury types (e.g., ACL tears, ankle sprains, hamstring strains) have distinct biomechanical mechanisms, risk factors, and intervention pathways. This limitation underscores the need for injury-specific prediction models and contextualizes both model performance and the interpretation of feature importance.

Model performance

Among the included studies, AUC was the most frequently reported metric for evaluating model performance, primarily reflecting the ability of a model to discriminate between positive and negative cases across varying thresholds. However, in the highly imbalanced context of lower limb injury prediction (average imbalance ratio = 0.35), a high AUC does not necessarily indicate satisfactory identification of the minority class, namely the actual injury cases (Van Eetvelde et al., 2021). To address this limitation, several studies additionally reported sensitivity and specificity to provide a more comprehensive assessment of clinical utility (Ayala et al., 2019; Javier Robles-Palazon et al., 2023; Kolodziej et al., 2023; Lopez-Valenciano et al., 2018; Oliver et al., 2020; Ruiz-Perez et al., 2021). In real-world sports injury prevention, practitioners often adopt a strategy of "erring on the side of caution"—prioritizing the identification of high-risk individuals even at the cost of increased false positives—thereby making higher sensitivity particularly important (Florkowski, 2008). Nevertheless, among the six studies in this review that reported both sensitivity and specificity, sensitivity values ranged from 0.35 to 0.78, whereas specificity ranged from 0.62 to 0.84. Notably, only one study demonstrated higher sensitivity than specificity (Ruiz-Perez et al., 2021), while the remaining studies showed the opposite pattern, including one with a sensitivity as low as (Kolodziej et al., 2023). From a clinical perspective, such imbalances indicate that many models are more effective at correctly identifying non-injury cases than detecting minority injury events, which may limit their utility for timely injury prevention and early intervention—settings where high sensitivity is particularly important.

While machine learning models generally demonstrate competitive predictive performance, they do not consistently outperform traditional statistical approaches. For example, Jauhiainen et al.(Jauhiainen et al., 2021) reported that LR achieved a slightly higher AUC (0.65) than RF (0.63) in youth athletes, and Oliver et al.(Oliver et al., 2020) similarly found LR (AUC = 0.69) to marginally exceed a DT model (AUC = 0.66) in elite youth soccer players. However, when evaluating performance beyond AUC, substantial differences emerged. In Oliver et al., the DT achieved markedly higher sensitivity (55.6%) compared with LR (11.1%), despite similar AUC values. This discrepancy highlights a critical issue: under class-imbalanced conditions common in injury datasets, AUC alone may mask models' ability to correctly identify injury cases. Thus, the mixed findings do not indicate a fundamental limitation of LR per se, but rather emphasize that model evaluation must account for metrics sensitive to minority-class detection when comparing ML with traditional methods.

In prospective injury prediction studies, class imbalance is a pervasive challenge, as injury cases are typically much less frequent than non-injury cases. Addressing this imbalance is therefore critical for robust model development. Among the 15 studies included, 10 (67%) applied specific strategies to handle imbalance, primarily resampling or cost-sensitive learning. Resampling was the most common (80%), involving techniques such as synthetic minority oversampling (e.g., SMOTE) to generate new "injury" samples, or under-sampling combined with ensemble learning (e.g., under-sampling bagging) to reduce "non-injury" samples. Although SMOTE was applied in six studies, three of them reported that its use did not improve predictive performance (Jauhiainen et al., 2022; Lopez-Valenciano et al., 2018; Ruddy et al., 2018). This pattern reflects a broader limitation of over-sampling in injury prediction: when synthetic samples are generated from nearest neighbors, the minority class may be overly homogenized, masking rare but clinically informative patterns and increasing overfitting risk (Carvalho et al., 2025; Fernández et al., 2018). In contrast, López-Valenciano et al. (2018) observed marginal gains using random under-sampling, which avoids synthetic noise but removes substantial majority-class information that may be essential for stable decision boundaries. Together, these findings illustrate a central methodological challenge in injury prediction: conventional resampling techniques often fail to capture the complex, low-prevalence nature of injury events. This suggests that using data-driven recommendation systems, such as those based on dataset complexity measures, to automatically identify the most appropriate resampling strategy may offer a more effective solution (Carvalho et al., 2025).

A frequently cited example is the study by Rommers et al. (2020), which prospectively monitored 734 elite youth soccer players (U10 - U15) across a full competitive season. Their models achieved balanced predictive performance (accuracy, sensitivity, and specificity all = 0.85) when forecasting both acute and overuse injuries. The study adopted a clear and standardized injury definition, recording any physical complaint that required evaluation by medical or paramedical staff; medical personnel were present at every training session and match, ensuring complete medical-attention reporting. Injuries included both event-related acute cases and overuse injuries without a single causal incident, and predictions covered injuries across the entire body rather than focusing on a specific anatomical region. Two factors likely contributed to the model's favorable performance. First, the dataset was unusually well balanced (50.1% injured vs. 49.9% non-injured), which helped minimize the class imbalance issues that typically challenge injury prediction models (Javier Robles-Palazon et al., 2023; Kolodziej et al., 2023; Ruiz-Perez et al., 2021). In a balanced dataset, ML models may be better positioned to learn injury-related patterns because the minority class is more adequately represented during training. Evidence from youth soccer injury-prediction studies suggests that models developed from relatively balanced class distributions (e.g., Rommers et al., 2020), IR ≈ 1.0) tend to report higher AUC values (≈ 0.85) compared with those trained on more imbalanced datasets (IR = 0.21-0.39, AUC = 0.66-0.70)(Javier Robles-Palazon et al., 2023; Oliver et al., 2020). While these findings do not establish a causal relationship, they indicate that class balance can contribute to improved predictive performance under certain conditions. Second, the adolescent sample (mean age = 11.7 ± 1.7 years) falls within a developmental period where injury risk shows clear age-related variation, with the 13 to 15 age range identified as the peak-incidence period (Rumpf and Cronin, 2012)—making such patterns easier for machine learning algorithms to detect (Jauhiainen et al., 2022). The study predicted injuries across the entire body, which may have improved overall model accuracy and stability. However, this broad classification reduces the ability to provide actionable guidance for specific anatomical sites. Predicting injuries by region would allow for more targeted prevention strategies and tailored interventions, which are typically more relevant in clinical practice.

Clinical translation challenges

In the clinical translation of ML for lower limb injury prediction, although some studies have reported strong model performance over extended prediction windows (AUC ≥ 0.8)(Ayala et al., 2019; Connaboy et al., 2019; Huang et al., 2023), their clinical applicability remains limited. These models frequently adopt long-term injury outcomes (e.g., across a season or a year) as labels. While this approach facilitates the accumulation of sufficient injury cases and mitigates the problem of "extreme class imbalance," it may compromise the temporal validity of predictions. On the one hand, athletes' risk status dynamically fluctuates with variations in training load and physiological condition (Bache-Mathiesen et al., 2022; Johnston et al., 2019). Because most injury-prediction studies in this field adopt a prospective design (Van Eetvelde et al., 2021), the predictor data are collected before the injury occurs. However, data obtained several months prior to the injury may still fail to reflect the athlete's immediate pre-injury condition. On the other hand, excessively long prediction windows reduce the actionable value of risk alerts, thereby constraining their utility for training monitoring and rehabilitation management. To enhance clinical feasibility, future research should investigate modeling strategies based on periodic screenings (e.g., monthly or per training cycle) to capture risk features closer to injury onset, thereby improving both the timeliness and practical relevance of predictions.

Notably, attempts have been made to develop short-term injury prediction models. For example, Briand et al. (Briand et al., 2022) proposed a framework for predicting injuries within 1-7 days, but its average sensitivity was only 0.35 ± 0.19, underscoring the methodological challenges associated with sample distribution and feature sensitivity in short-term predictions. More recently, a four-year longitudinal study in professional football applied machine learning to internal (RPE) and external (GPS-derived) workload data from the two-week and four-week periods prior to injury (Martins et al., 2025). Using a four-week window, the KStar classifier achieved a sensitivity of 0.69, a specificity of 0.76, and an AUC of 0.81. The two-week models delivered slightly lower but still meaningful predictive performance, with the MLP yielding a sensitivity of 0.75, a specificity of 0.69, and an AUC of 0.79.Collectively, these findings further demonstrate that short-term injury risk prediction is achievable when leveraging multidimensional workload indicators.

A further challenge that directly influences clinical translation, yet is often overlooked in existing reviews, is the heterogeneity of injury mechanisms included in model development. Previous syntheses (Leckey et al., 2025; Van Eetvelde et al., 2021; Yuan et al., 2025) did not systematically distinguish between contact and non-contact injuries in their inclusion criteria. This lack of differentiation leads to pooled evidence combining fundamentally different etiological pathways: contact injuries are frequently driven by external forces or collisions, whereas non-contact injuries are more closely linked to intrinsic factors, neuromuscular control, and biomechanical patterns (Dauty et al., 2022; Yu and Garrett, 2007). Aggregating these mechanisms may obscure true model performance, alter feature importance profiles, and reduce the generalizability of findings. By contrast, the present review adopts a strictly defined non-contact lower-limb injury criterion, reducing etiological heterogeneity and enabling a more coherent evaluation of prediction models within a mechanistically consistent category. This focus provides clearer insight into which data modalities, feature representations, and ML architectures are effective for non-contact injury risk and strengthens the translational relevance of the synthesized evidence.

Looking forward, emerging methodological frameworks provide promising avenues for improving model timeliness and contextual relevance. The Weighted Cumulative Exposure (WCE) approach, implemented within Piecewise Exponential Additive Mixed Models allows researchers to model how past training loads accumulate and exert time-dependent effects on injury risk (Zumeta-Olaskoaga et al., 2025). These models flexibly estimate the time window during which previous exposures meaningfully contribute to current injury hazard, enabling predictions that better reflect the evolving load patterns experienced in real-world training environments.

Interpretability

In the field of lower limb injury prediction, ML models have demonstrated promising predictive performance; however, their practical utility extends beyond conventional metrics such as accuracy or AUC. A critical issue is whether these models can be reliably trusted in clinical or sports settings. Trustworthiness depends not only on predictive capability but also on model interpretability and reliability. Cross-validation plays an essential role in this context, providing a more robust estimate of model generalizability and reducing the risk of overfitting. Notably, all 15 studies included in this review employed cross-validation procedures, underscoring its role as a standard methodological safeguard. Nevertheless, cross-validation alone does not guarantee clinical or applied reliability. Conventional cross-validation can produce overly optimistic performance estimates when data exhibit temporal dependence, as is common in training-load-based injury prediction (Roberts et al., 2017). This highlights the need for time-aware validation strategies and, more broadly, for external validation on independent cohorts. Complementing cross-validation with external validation and domain-relevant interpretability is therefore essential to ensure real-world trustworthiness(Ramspek et al., 2021). Ultimately, the goal of injury prediction is not only to identify high-risk individuals but also to reveal actionable mechanisms underlying injury risk. Analogous to the established link between smoking and cancer, interpretable models can inform targeted intervention strategies (Wang et al., 1999). Consequently, interpretability constitutes a key prerequisite for translating ML models from research into practice.

Existing literature shows considerable variability in how interpretability is conceptualized within ML-based injury prediction. Prior reviews, such as those by Leckey et al. (2025) and Yuan et al. (2025), have largely centered their discussion on post-hoc explanation techniques, particularly SHAP, to interpret complex “black-box” models including XGBoost, neural networks and SVM. While these methods are valuable for quantifying feature contributions, they represent only one dimension of model interpretability. In contrast, the present review underscores the importance of inherently interpretable “white-box” models such as DT and RF. These algorithms offer transparency by design, enabling direct inspection of decision pathways and feature relevance without external interpretability tools (Belle and Papantonis, 2021). This is one reason why tree-based models remain prevalent in injury-prediction research, as their structure supports accessible metrics of feature importance (for example, split frequency or impurity-based measures) that facilitate clear identification of salient risk factors. Empirical studies further illustrate the advantages of these models. López-Valenciano et al. (2018) used DT classifiers to highlight previous injury history and strength asymmetries as primary determinants of lower-extremity injury risk. Similarly, Ruiz-Pérez et al. (2021) applied RF and identified workload and neuromuscular parameters as dominant predictors based on impurity-based importance scores. These examples demonstrate how tree-based approaches not only reveal influential variables but also clarify how these features interact to stratify athletes into different risk profiles. Such transparency is particularly valuable in applied sport settings where practitioners must interpret and justify risk assessments.

Compared with the SHAP-centric approach in previous syntheses, our broader framing highlights that model interpretability can arise either from intrinsic model structure or from post-hoc explanation techniques applied to more complex architectures. Recognizing both pathways provides a more comprehensive understanding of how ML outputs can inform mechanism-oriented interpretations and guide evidence-based intervention design (Kulshrestha et al., 2021; Majumdar et al., 2022).

At the feature level, several studies have identified relatively stable risk factors. When demographic variables were included, BMI and previous injury history frequently emerged as key predictors, consistent with broader musculoskeletal injury literature (Hecksteden et al., 2023; Rommers et al., 2020). In addition to commonly used demographic and biomechanical variables, many studies have incorporated psychological measures into their models (Ayala et al., 2019; Javier Robles-Palazon et al., 2023; Lopez-Valenciano et al., 2018; Ruiz-Perez et al., 2021). Notably, Lipps Lene et al. (2024) directly compared models with and without psychological factors and found that adding these variables significantly improved predictive performance ($p < 0.001$).

However, evidence across studies also shows considerable variability in the relative importance of individual predictors, which complicates their clinical use. Ruddy et al. (2018) examined whether supervised learning models using preseason eccentric hamstring strength, age, and previous HSI history could accurately predict hamstring strain injuries in elite Australian footballers. Although the models were trained on the same dataset, performance fluctuated widely (AUC 0.24-0.92) due to minor changes in training-testing partitions. This instability reflected meaningful season-to-season differences in cohort characteristics: injured players were substantially weaker than uninjured players in 2013, whereas no strength differences were observed in 2015 despite similar HSI incidence. These findings demonstrate that the influence of commonly cited risk factors is highly context dependent and shaped by variations in conditioning status, training load, and population profiles. Consequently, predictors identified in one season or team may not generalize reliably to others.

Further evidence of contextual fluctuation is provided by Ayala et al. (2019), who integrated neuromuscular, personal, and psychological variables into an injury-specific model. They observed that no single predictor consistently dominated across classifiers. Instead, variables such as sleep quality, hip flexion range of motion, and angle-specific torque contributed variably, reflecting the inherently multifactorial nature of HSI etiology. Importantly, their injury-specific modeling strategy produced stronger predictive performance than studies relying on limited or non-specific feature sets, suggesting that predictor stability improves when models are grounded in mechanisms directly relevant to the injury being predicted.

Taken together, these findings indicate that the relevance of individual predictors varies substantially across seasons, populations, and modeling frameworks. Therefore, machine-learning-derived predictors should not be assumed to generalize across contexts unless they are rooted in injury-specific mechanisms and validated across multiple cohorts. For clinical application, this underscores the importance of developing models that incorporate comprehensive, injury-relevant features and that undergo external validation before being used to guide risk-mitigation strategies.

Limitations

Despite systematically reviewing current advances in applying ML to lower limb injury prediction, several limitations should be acknowledged. First, although all included studies used some form of internal cross-validation such as k-fold or leave-one-out, considerable methodological heterogeneity remained across studies in terms of study populations, injury types, feature engineering strategies, and prediction windows. More importantly, most studies relied only on internal validation and did not conduct independent external validation, which limits the generalizability of model performance and may contribute to inconsistencies in the reported findings. Second, the transparency of ML methodology in the included studies was limited. Many studies provided insufficient detail regarding model development pipelines, hyperparameter tuning procedures, software toolboxes, and code availability. Differences in how model interpretability was conceptualized and implemented, together with variation in injury sites studied, further hinder cross-study comparisons and reduce the feasibility of systematic integration and clinical translation. Third, this review included only peer-reviewed publications written in English, excluding non-English articles, theses, conference papers, and grey literature. Although this approach enhances methodological rigor, it may also have resulted in the omission of relevant evidence. Finally, this review synthesized findings qualitatively and did not perform a meta-analysis. The absence of pooled effect estimates prevents direct quantitative comparisons of ML algorithm performance. Therefore, the findings should be interpreted cautiously, and future research, especially large-scale multicenter studies with transparent methodological reporting and external validation, is needed to strengthen and extend these conclusions.

Conclusion

This review demonstrates that ML holds considerable potential for predicting non-contact lower limb injuries; however, its clinical utility depends not only on predictive performance but also on interpretability and reliability. White-box algorithms offer inherent transparency, enhancing clinical comprehensibility, whereas black-box models, despite achieving higher predictive accuracy, face limitations in trustworthiness due to their opacity. Therefore, future research should strive to balance predictive performance with interpretability by integrating post-hoc explanation techniques and hybrid modeling frameworks to facilitate clinical translation. Moreover, standardized data collection and feature selection, integration of multi-source information, and large-scale prospective studies are critical for enhancing model robustness and generalizability across populations. Overall, only through the coordinated development of predictive performance, interpretability, and methodological rigor can ML truly support precision injury prevention and rehabilitation in sports practice.

Acknowledgements

This work was supported by Key Project of Humanities and Social Sciences in Anhui Province Universities (2023AH050883, 2024AH052247); Major Project of Philosophy and Social Sciences in Anhui Province Universities (2023AH040116). The authors declare that there are no conflicts of interest. The experiments comply with the current laws of the country where they were performed. The data that support the findings of this study are available on request from the corresponding author.

References

Achenbach, L., Klein, C., Luig, P., Bloch, H., Schneider, D. and Fehske, K. (2021) Collision with opponents—but not foul play—dominates injury mechanism in professional men's basketball. *BMC Sports Science, Medicine and Rehabilitation* **13**, 94. https://doi.org/10.1186/s13102-021-00322-z

Al Attar, W. S. A., Soomro, N., Sinclair, P. J., Pappas, E. and Sanders, R. H. (2017) Effect of injury prevention programs that include the Nordic hamstring exercise on hamstring injury rates in soccer players: a systematic review and meta-analysis. *Sports Medicine* **47**, 907-916. https://doi.org/10.1007/s40279-016-0638-2

Ayala, F., López-Valenciano, A., Martín, J. A. G., Croix, M. D. S., Vera-Garcia, F. J., del Pilar García-Vaquero, M., Ruiz-Pérez, I. and Myer, G. D. (2019) A preventive model for hamstring injuries in professional soccer: Learning algorithms. *International Journal of Sports Medicine* **40**, 344-353. https://doi.org/10.1055/a-0826-1955

Bache-Mathiesen, L. K., Andersen, T. E., Dalen-Lorentsen, T., Clarsen, B. and Fagerland, M. W. (2022) Assessing the cumulative effect of long-term training load on the risk of injury in team sports. *BMJ Open Sport & Exercise Medicine* **8**, e001342. https://doi.org/10.1136/bmjsem-2022-001342

Bartlett, J. D., O'Connor, F., Pitchford, N., Torres-Ronda, L. and Robertson, S. J. (2017) Relationships between internal and external training load in team-sport athletes: evidence for an individualized approach. *International Journal of Sports Physiology and Performance* **12**, 230-234. https://doi.org/10.1123/ijspp.2015-0791

Belkhelladi, M., Cierson, T. and Martineau, P. A. (2025) Biomechanical Risk Factors for Increased Anterior Cruciate Ligament Loading and Injury: A Systematic Review. *Orthopaedic Journal of Sports Medicine* **13**, 23259671241312681. https://doi.org/10.1177/23259671241312681

Belle, V. and Papantonis, I. (2021) Principles and practice of explainable machine learning. *Frontiers in Big Data* **4**, 688969. https://doi.org/10.3389/fdata.2021.688969

Bittencourt, N. F., Meeuwisse, W., Mendonça, L., Nettel-Aguirre, A., Ocarino, J. and Fonseca, S. (2016) 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* **50**, 1309-1314. https://doi.org/10.1136/bjsports-2015-095850

Bogaert, S., Davis, J., Van Rossom, S. and Vanwanseele, B. (2022) Impact of Gender and Feature Set on Machine-Learning-Based Prediction of Lower-Limb Overuse Injuries Using a Single Trunk-Mounted Accelerometer. *Sensors (Basel)* **22**, 2874. https://doi.org/10.3390/s22082874

Briand, J., Deguire, S., Gaudet, S. and Bieuzen, F. (2022) Monitoring variables influence on random forest models to forecast injuries in short-track speed skating. *Frontiers in Sports and Active Living* **4**, 896828. https://doi.org/10.3389/fsals.2022.896828

Carvalho, M., Pinho, A. J. and Brás, S. (2025) Resampling approaches to handle class imbalance: a review from a data perspective. *Journal of Big Data* **12**, 71. https://doi.org/10.1186/s40537-025-01119-4

Chia, L., De Oliveira Silva, D., Whalan, M., McKay, M. J., Sullivan, J., Fuller, C. W. and Pappas, E. (2022) Non-contact anterior cruciate ligament injury epidemiology in team-ball sports: a systematic review with meta-analysis by sex, age, sport, participation level, and exposure type. *Sports Medicine* **52**, 2447-2467. https://doi.org/10.1007/s40279-022-01697-w

Connaboy, C., Eagle, S. R., Johnson, C. D., Flanagan, S. D., Mi, Q. and Nindl, B. C. (2019) Using Machine Learning to Predict Lower-Extremity Injury in US Special Forces. *Medicine and Science in Sports and Exercise* **51**, 1073-1079. https://doi.org/10.1249/MSS.0000000000001900

Dauty, M., Crenn, V., Louguet, B., Grondin, J., Menu, P. and Fouasson-Chailloux, A. (2022) Anatomical and neuromuscular factors associated to non-contact anterior cruciate ligament injury. *Journal of Clinical Medicine* **11**, 1402. https://doi.org/10.3390/jcm11051402

Ekstrand, J., Hägglund, M. and Waldén, M. (2011) Epidemiology of muscle injuries in professional football (soccer). *The American Journal of Sports Medicine* **39**, 1226-1232. https://doi.org/10.1177/0363546510395879

Evans, S. L., Owen, R., Whittaker, G., Davis, O. E., Jones, E. S., Hardy, J. and Owen, J. (2024) Non-contact lower limb injuries in Rugby Union: A two-year pattern recognition analysis of injury risk factors. *Plos One* **19**, e0307287. https://doi.org/10.1371/journal.pone.0307287

Fernández, A., Garcia, S., Herrera, F. and Chawla, N. V. (2018) SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research* **61**, 863-905. https://doi.org/10.1613/jair.1.11192

Fernandez-Felix, B. M., López-Alcalde, J., Roqué, M., Muriel, A. and Zamora, J. (2023) CHARMS and PROBAST at your fingertips: a template for data extraction and risk of bias assessment in systematic reviews of predictive models. *BMC Medical Research Methodology* **23**, 44. https://doi.org/10.1186/s12874-023-01849-0

Florkowski, C. M. (2008) Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *The Clinical Biochemist Reviews* **29**, S83-S87.

Green, B., Bourne, M. N., Van Dyk, N. and Pizzari, T. (2020) Recalibrating the risk of hamstring strain injury (HSI): A 2020 systematic review and meta-analysis of risk factors for index and recurrent hamstring strain injury in sport. *British Journal of Sports Medicine* **54**, 1081-1088. https://doi.org/10.1136/bjsports-2019-100983

Guan, Y., Bredin, S. S., Taunton, J., Jiang, Q., Wu, N., Li, Y. and Warburton, D. E. (2021) Risk factors for non-contact lower-limb injury: a retrospective survey in pediatric-age athletes. *Journal of Clinical Medicine* **10**, 3171. https://doi.org/10.3390/jcm10143171

Hecksteden, A., Schmartz, G. P., Egyptien, Y., Aus der Fünten, K., Keller, A. and Meyer, T. (2023) Forecasting football injuries by combining screening, monitoring and machine learning. *Science and Medicine in Football* **7**, 214-228. https://doi.org/10.1080/24733938.2022.2095006

Henriquez, M., Sumner, J., Faherty, M., Sell, T. and Bent, B. (2020) Machine learning to predict lower extremity musculoskeletal injury risk in student athletes. *Frontiers in Sports and Active Living* **2**, 576655. https://doi.org/10.3389/fsals.2020.576655

Hickey, J., Shield, A. J., Williams, M. D. and Opar, D. A. (2014) The financial cost of hamstring strain injuries in the Australian Football League. *British Journal of Sports Medicine* **48**, 729-730. https://doi.org/10.1136/bjsports-2013-092234

Horvat, T. and Job, J. (2020) The use of machine learning in sport outcome prediction: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **10**, e1380. https://doi.org/10.1002/widm.1380

Huang, Y., Huang, S., Wang, Y., Li, Y., Gui, Y. and Huang, C. (2022) A novel lower extremity non-contact injury risk prediction model based on multimodal fusion and interpretable machine learning. *Frontiers in Physiology* **13**, 1024286. https://doi.org/10.3389/fphys.2022.937546

Huang, Y., Li, C., Bai, Z., Wang, Y., Ye, X., Gui, Y. and Lu, Q. (2023) The impact of sport-specific physical fitness change patterns on lower limb non-contact injury risk in youth female basketball players: a pilot study based on field testing and machine learning. *Frontiers in Physiology* **14**, 1176713.

https://doi.org/10.3389/fphys.2023.1182755

Hubáček, O., Šourek, G. and Železný, F. (2019) Learning to predict soccer results from relational data with gradient boosted trees. *Machine Learning* **108**, 29-47. https://doi.org/10.1007/s10994-018-5704-6

Jauhiainen, S., Kauppi, J.-P., Krosshaug, T., Bahr, R., Bartsch, J. and Äyrämö, S. (2022) Predicting ACL injury using machine learning on data from an extensive screening test battery of 880 female elite athletes. *The American Journal of Sports Medicine* **50**, 2917-2924. https://doi.org/10.1177/03635465221112095

Jauhiainen, S., Kauppi, J.-P., Leppänen, M., Pasanen, K., Parkkari, J., Vasankari, T., Kannus, P. and Äyrämö, S. (2021) New machine learning approach for detection of injury risk factors in young team sport athletes. *International Journal of Sports Medicine* **42**, 175-182. https://doi.org/10.1055/a-1231-5304

Robles-Palazon, F. J., Puerta-Callejon, J. M., Gamez, J. A., Croix, M. D. S., Cejudo, A., Santonja, F., de Baranda, P. S. and Ayala, F. (2023) Predicting injury risk using machine learning in male youth soccer players. *Chaos, Solitons & Fractals* **167**, 113062. https://doi.org/10.1016/j.chaos.2022.113079

Johnston, R., Cahalan, R., Bonnett, L., Maguire, M., Nevill, A., Glasgow, P., O'Sullivan, K. and Comyns, T. (2019) Training load and baseline characteristics associated with new injury/pain within an endurance sporting population: a prospective study. *International Journal of Sports Physiology and Performance* **14**, 590-597. https://doi.org/10.1123/ijspp.2018-0644

Kolodziej, M., Groll, A., Nolte, K., Willwacher, S., Alt, T., Schmidt, M. and Jaitner, T. (2023) Predictive modeling of lower extremity injury risk in male elite youth soccer players using least absolute shrinkage and selection operator regression. *Scandinavian Journal of Medicine & Science in Sports* **33**, 1021-1033. https://doi.org/10.1111/sms.14312

Kulshrestha, S., Dligach, D., Joyce, C., Gonzalez, R., O'Rourke, A. P., Glazer, J. M., Stey, A., Kruser, J. M., Churpek, M. M. and Afshar, M. (2021) Comparison and interpretability of machine learning models to predict severity of chest injury. *JAMIA Open* **4**, ooab015. https://doi.org/10.1093/jamiaopen/ooab015

Leckey, C., Van Dyk, N., Doherty, C., Lawlor, A. and Delahunt, E. (2025) Machine learning approaches to injury risk prediction in sport: a scoping review with evidence synthesis. *British Journal of Sports Medicine* **59**, 491-500. https://doi.org/10.1136/bjsports-2024-108576

Lipps Lene, C., Frere, J. and Weissland, T. (2024) Machine learning in knee injury sequelae detection: Unravelling the role of psychological factors and preventing long-term sequelae. *Journal of Experimental Orthopaedics* **11**, e70081. https://doi.org/10.1002/jeo2.70081

Liveris, N. I. (2025) Applying systems thinking approaches to investigate the complex interrelationships of risk factors affecting acute non-contact lower limb injuries in team sports (PhD Academy Award). *British Journal of Sports Medicine* **59**, 683-684. https://doi.org/10.1136/bjsports-2025-109742

López-Valenciano, A., Ayala, F., Puerta, J. M., De Ste Croix, M. B. A., Vera-Garcia, F. J., Hernandez-Sanchez, S., Ruiz-Perez, I. and Myer, G. D. (2018) A preventive model for muscle injuries: a novel approach based on learning algorithms. *Medicine and Science in Sports and Exercise* **50**, 915-927. https://doi.org/10.1249/MSS.0000000000001523

López-Valenciano, A., Ruiz-Pérez, I., Garcia-Gómez, A., Vera-Garcia, F. J., Croix, M. D. S., Myer, G. D. and Ayala, F. (2020) Epidemiology of injuries in professional football: a systematic review and meta-analysis. *British Journal of Sports Medicine* **54**, 711-718. https://doi.org/10.1136/bjsports-2018-099577

Lu, D., McCall, A., Jones, M., Steinweg, J., Gelis, L., Fransen, J. and Duffield, R. (2021) The financial and performance cost of injuries to teams in Australian professional soccer. *Journal of Science and Medicine in Sport* **24**, 463-467. https://doi.org/10.1016/j.jsams.2020.10.013

Lu, Y., Pareek, A., Lavoie-Gagne, O. Z., Forlenza, E. M., Patel, B. H., Reinholz, A. K., Forsythe, B. and Camp, C. L. (2022) Machine learning for predicting lower extremity muscle strain in National Basketball Association athletes. *Orthopaedic Journal of Sports Medicine* **10**. https://doi.org/10.1177/23259671221111742

Majumdar, A., Bakirov, R., Hodges, D., Scott, S. and Rees, T. (2022) Machine learning for understanding and predicting injuries in football. *Sports Medicine - Open* **8**, 73. https://doi.org/10.1186/s40798-022-00465-4

Martins, F., Sarmento, H., Gouveia, É. R., Saveca, P. and Przednowek, K. (2025) Machine learning-based prediction of muscle injury risk in professional football: a four-year longitudinal study. *Journal of Clinical Medicine* **14**, 8039. https://doi.org/10.3390/jcm14138039

Musat, C. L., Mereuta, C., Nechita, A., Tutunaru, D., Voipan, A. E., Voipan, D., Mereuta, E., Gurau, T. V., Gurău, G. and Nechita, L. C. (2024) Diagnostic applications of AI in sports: a comprehensive review of injury risk prediction methods. *Diagnostics* **14**, 2516. https://doi.org/10.3390/diagnostics140202516

Nassis, G., Verhagen, E., Brito, J., Figueiredo, P. and Krustrup, P. (2023) A review of machine learning applications in soccer with an emphasis on injury risk. *Biology of Sport* **40**, 233-239. https://doi.org/10.5114/biolsport.2023.114283

Oliver, J. L., Ayala, F., De Ste Croix, M. B., Lloyd, R. S., Myer, G. D. and Read, P. J. (2020) Using machine learning to improve our understanding of injury risk and prediction in elite male youth football players. *Journal of Science and Medicine in Sport* **23**, 1044-1048. https://doi.org/10.1016/j.jsams.2020.03.006

Ou-Yang, Y., Hong, W., Peng, L., Mao, C.-X., Zhou, W.-J., Zheng, W.-T., Wang, Q., Qi, F., Li, X.-W. and Chen, S.-H. (2025) Explaining basketball game performance with SHAP: insights from Chinese Basketball Association. *Scientific Reports* **15**, 13793. https://doi.org/10.1038/s41598-025-97817-3

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A. and Brennan, S. E. (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71. https://doi.org/10.1136/bmj.n71

Ramspek, C. L., Jager, K. J., Dekker, F. W., Zoccali, C. and van Diepen, M. (2021) External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal* **14**, 49-58. https://doi.org/10.1093/ckj/sfz143

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B. and Thuiller, W. (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913-929. https://doi.org/10.1111/ecog.02881

Rommers, N., Rössler, R., Verhagen, E., Vandecasteele, F., Verstockt, S., Vaeyens, R., Lenoir, M., D'hondt, E. and Witvrouw, E. (2020) A machine learning approach to assess injury risk in elite youth football players. *Medicine and Science in Sports and Exercise* **52**, 1745-1751. https://doi.org/10.1249/MSS.0000000000002304

Rossi, A., Pappalardo, L., Cintia, P., Iaia, F. M., Fernández, J. and Medina, D. (2018) Effective injury forecasting in soccer with GPS training data and machine learning. *PLOS ONE* **13**, e0201264. https://doi.org/10.1371/journal.pone.0201264

Rössler, R., Junge, A., Bizzini, M., Verhagen, E., Chomiak, J., aus der Fünten, K., Meyer, T., Dvorak, J., Lichtenstein, E. and Beaudouin, F. (2018) A multinational cluster randomised controlled trial to assess the efficacy of '11+ Kids': a warm-up programme to prevent injuries in children's football. *Sports Medicine* **48**, 1493-1504. https://doi.org/10.1007/s40279-017-0834-8

Ruddy, J. D., Cormack, S. J., Whiteley, R., Williams, M. D., Timmins, R. G. and Opar, D. A. (2019) Modeling the risk of team sport injuries: a narrative review of different statistical approaches. *Frontiers in Physiology* **10**, 829. https://doi.org/10.3389/fphys.2019.00829

Ruddy, J. D., Shield, A. J., Maniar, N., Williams, M. D., Duhig, S. J., Timmins, R. G., Hickey, J., Bourne, M. N. and Opar, D. A. (2018) Predictive modeling of hamstring strain injuries in elite Australian footballers. *Medicine and Science in Sports and Exercise* **50**, 906-914. https://doi.org/10.1249/MSS.0000000000001524

Ruiz-Perez, I., Lopez-Valenciano, A., Hernandez-Sanchez, S., Puerta-Callejon, J. M., De Ste Croix, M., Sainz de Baranda, P. and Ayala, F. (2021) A field-based approach to determine soft tissue injury risk in elite futsal using novel machine learning techniques. *Frontiers in Psychology* **12**, 610210. https://doi.org/10.3389/fpsyg.2021.610210

Rumpf, M. C. and Cronin, J. (2012) Injury incidence, body site, and severity in soccer players aged 6-18 years: implications for injury prevention. *Strength and Conditioning Journal* **34**, 20-31. https://doi.org/10.1519/SSC.0b013e31821a9833

Sampaio, T., Oliveira, J. P., Marinho, D. A., Neiva, H. P. and Morais, J.

E. (2024) Applications of machine learning to optimize tennis performance: a systematic review. *Applied Sciences* **14**, 5517. https://doi.org/10.3390/app14085517

Van Eetvelde, H., Mendonça, L. D., Ley, C., Seil, R. and Tischer, T. (2021) Machine learning methods in sport injury prediction and prevention: a systematic review. *Journal of Experimental Orthopaedics* **8**, 1-15. https://doi.org/10.1186/s40634-021-00346-x

Wang, H.-X., Fratiglioni, L., Frisoni, G. B., Viitanen, M. and Winblad, B. (1999) Smoking and the occurrence of Alzheimer's disease: cross-sectional and longitudinal data in a population-based study. *American Journal of Epidemiology* **149**, 640-644. https://doi.org/10.1093/oxfordjournals.aje.a009866

Watson, N., Hendricks, S., Stewart, T. and Durbach, I. (2021) Integrating machine learning and decision support in tactical decision-making in rugby union. *Journal of the Operational Research Society* **72**, 2274-2285. https://doi.org/10.1080/01605682.2020.1779624

Webster, K. E. and Hewett, T. E. (2018) Meta-analysis of meta-analyses of anterior cruciate ligament injury reduction training programs. *Journal of Orthopaedic Research* **36**, 2696-2708. https://doi.org/10.1002/jor.24043

Whittaker, J. L., Räisänen, A. M., Martin, C., Galarneau, J.-M., Martin, M., Losciale, J. M., Bullock, G. S., Dubé, M.-O., Bizzini, M. and Bourne, M. N. (2025) Modifiable risk factors for lower-extremity injury: a systematic review and meta-analysis for the Female, Woman and Girl Athlete Injury Prevention (FAIR) consensus. *British Journal of Sports Medicine.* https://doi.org/10.1016/j.jsams.2025.07.096

Wilkerson, G. B., Gupta, A. and Colston, M. A. (2018) Mitigating sports injury risks using internet of things and analytics approaches. *Risk Analysis* **38**, 1348-1360. https://doi.org/10.1111/risa.12932

Willy, R. W. (2018) Innovations and pitfalls in the use of wearable devices in the prevention and rehabilitation of running related injuries. *Physical Therapy in Sport* **29**, 26-33. https://doi.org/10.1016/j.ptsp.2017.10.003

Wolff, R. F., Moons, K. G., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., Mallett, S. and PROBAST Group. (2019) PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine* **170**, 51-58. https://doi.org/10.7326/M18-1376

Yu, B. and Garrett, W. E. (2007) Mechanisms of non-contact ACL injuries. *British Journal of Sports Medicine* **41**, i47-i51. https://doi.org/10.1136/bjsm.2007.037192

Yuan, J., Zeng, Q., Li, J., Cong, Z. and Zhang, Y. (2025) Machine learning applications in sports injury prediction: a narrative review. *Science Progress* **108**, 00368504251385956. https://doi.org/10.1177/00368504251385956

Zumeta-Olaskoaga, L., Bender, A. and Lee, D.-J. (2025) Flexible modelling of time-varying exposures and recurrent events to analyse training load effects in team sports injuries. *Journal of the Royal Statistical Society Series C: Applied Statistics* **74**, 391-405. https://doi.org/10.1093/jrsssc/qlae059

Key points

- Tree-based ML algorithms dominate non-contact lower limb injury prediction and generally demonstrate acceptable discriminative performance, yet sole reliance on AUC risks overlooking poor recognition in imbalanced datasets.
- Clinical translation faces challenges of long prediction windows, generalized injury types, and imbalance; short-term, specific, multi-source modelling may improve utility.
- Interpretability remains key for ML adoption; despite advances with white-box and post-hoc methods, heterogeneity highlights the need for standardized, mechanism-driven approaches.

AUTHOR BIOGRAPHY

JIN YUAN
Employment
School of Physical Education, Anhui Polytechnic University
Degree
Med
Research interests
Health promotion and machine learning, etc.
E-mail: 2231212128@stu.ahpu.edu.cn

QUANWEN ZENG
Employment
School of Physical Education, Anhui Polytechnic University
Degree
Med
Research interests
Physical Education and Training, etc.
E-mail: 2231212133@stu.ahpu.edu.cn

ANJIE WANG
Employment
School of Physical Education, Anhui Polytechnic University
Degree
PhD
Research interests
Exercise physiology and performance, etc.
E-mail: wanganjie@ahpu.edu.cn

YONG ZHANG
Employment
School of Physical Education, Anhui Polytechnic University
Degree
PhD
Research interests
Exercise intervention and health promotion, etc.
E-mail: zhangyong@ahpu.edu.cn

JUN LI
Employment
School of Athletic Performance, Shanghai University of Sport
Degree
PhD
Research interests
Exercise physiology and health promotion, etc.
E-mail: lijun198112180978@126.com

✉ **Yong Zhang**
School of Physical Education, Anhui Polytechnic University, Wuhu, 241000, Anhui, China

✉ **Jun Li**
School of Athletic Performance, Shanghai University of Sport, Shanghai, 200438, China

Supplementary Materials

Supplementary Table 1. Search query.

PubMed
(((((((("Athletic Injuries"[Mesh]) OR (Injuries, Athletic)) OR (Athletic Injury)) OR (Injury, Athletic)) OR (Injuries, Sports)) OR (Injury, Sports)) OR (Sports Injury)) OR (Sports Injuries)) AND (((("Machine Learning"[Mesh]) OR (Learning, Machine [Title/Abstract])) OR (Transfer Learning[Title/Abstract])) OR (Learning, Transfer[Title/Abstract])) AND (((((((lower extremity[MeSH Terms]) OR (Extremities, Lower[Title/Abstract])) OR (Lower Extremities[Title/Abstract])) OR (Lower Limb[Title/Abstract])) OR (Limb, Lower[Title/Abstract])) OR (Limbs, Lower[Title/Abstract])) OR (Lower Limbs[Title/Abstract])) OR (Membrum inferius[Title/Abstract])
Web of science
((((((((TS=(Lower Extremity)) OR TS=(Extremities, Lower)) OR TS=(Lower Extremities)) OR TS=(Extremity, Lower)) OR TS=(Lower Limb)) OR TS=(Limb, Lower)) OR TS=(Limbs, Lower)) OR TS=(Lower Limbs)) OR TS=(Membrum inferius) AND (((TS=(Machine Learning)) OR TS=(Learning, Machine)) OR TS=(Transfer Learning)) OR TS=(Learning, Transfer) AND (((((((TS=(Athletic Injuries)) OR TS=(Injuries, Athletic)) OR TS=(Athletic Injury)) OR TS=(Injury, Athletic)) OR TS=(Injuries, Sports)) OR TS=(Injury, Sports)) OR TS=(Sports Injury)) OR TS=(Sports Injuries)
SPORTDiscus (EBSCO)
('Machine Learning' OR 'Learning, Machine' OR 'Transfer Learning' OR 'Learning, Transfer') AND ('Lower Extremity' OR 'Extremities, Lower' OR 'Lower Extremities' OR 'Extremity, Lower' OR 'Lower Limb' OR 'Limb, Lower' OR 'Limbs, Lower' OR 'Lower Limbs' OR 'Membrum inferius') AND ('Athletic Injuries' OR 'Injuries, Athletic' OR 'Athletic Injury' OR 'Injury, Athletic' OR 'Injuries, Sports' OR 'Injury, Sports' OR 'Sports Injury' OR 'Sports Injuries')

Supplementary Table 2. A summary of each reviewed paper.

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Lopez-Valenciano et al., 2018	Soccer, Handball	132(29)	Lower Extremity Muscle	DT	Age group, History of MUSINJ last season, Maximal level of play achieved; BMI, Sleep Quality, Sport Devaluation, YBalance, ISOM Hip Strength, Leg ROM, Core-USNF, Core-USWF, Core-USCD, Isokinetic Knee Muscle Torques, (MUSIN: Muscle injury; BMI: body mass index; Bila: bilateral; Uni: unilateral; ISOM. Isometric; Add: adduction; Abd: abduction; ROM: range of motion; ADF: ankle dorsi-flexion; KE: knee extension; KF: knee flexion; HF: hip flexion; APT: angle of peak torque; ECC: eccentric; CON: concentric; PT: peak torque; s: seconds; °: degree; USNF: unstable sitting without feedback; USWF: unstable sitting with feedback; USCD: unstable sitting while performing circular displacements with feedback)	AUC = 0.75, Sensitivity = 0.66, Specificity = 0.69
Ruddy et al., 2018	Football	362(53)	Hamstring Strain Injuries	Naïve Bayes	Age, Height, Mass, Playing position, ACL injury history, Hamstring strain 12-month injury history, Eccentric hamstring strength - peak force	Median AUC 2013 (Year): 0.60 2013 – 2015 (Year): 0.52 2015 (Year): 0.57
Ayala et al., 2019	Soccer	96(18)	Hamstring Strain Injuries	DT	Age, History of HSI last season, Maximal level of play achieved, Sleep quality, Physical/emotional exhaustion, Reduced sense of accomplishment, YBalance, PTISOM, ROM, Isokinetic leg strength/torque, (HSI: hamstring strain injury; Isometric hip, ankle and knee strength; ROM: range of motion; ADF: ankle dorsi-flexion; Q: quadriceps; H: hamstring; HF: hip flexion; HER: hip external rotation; Ant: anterior; Post: posterior; APT: angle of peak torque; ECC: eccentric; CON: concentric; PT: peak torque; T: torque; FUNC: functional; CONV: conventional; USNF: unstable sitting without feedback)	AUC = 0.84, Sensitivity = 0.78, Specificity = 0.84
Connaboy et al., 2019	Military personnel	140(38)	Lower Extremity	DT	Age, Height, Weight, Body fat, Aerobic capacity, Peak anaerobic power, Mean anaerobic power, R knee extension, R knee flexion, L knee extension, L knee flexion, R knee flexion/extension ratio, L knee flexion/extension ratio, R ankle eversion, L ankle eversion, R ankle inversion, L ankle inversion, R ankle eversion: inversion ratio, L ankle eversion/inversion ratio, Knee extension, Knee flexion, Ankle eversion, Ankle inversion, R hip extension, L hip extension, R knee active extension, L knee active extension, R ankle dorsiflexion, L ankle dorsiflexion, Hip extension, Knee active extension, Ankle dorsiflexion	AUC = 0.91

Supplementary Table 2. Continue…

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Oliver et al., 2020	Soccer	355(99)	Lower Extremity	DT	Age (y), Height (cm), Mass (kg), BMI (kg/m²), Leg Length (cm), Maturity-Offset, 75%Hop L PVGRF (BW), 75%Hop R PVGRF (BW), 75%Hop Asym (%), SLCMJ L PVGRF (BW), SLCMJ R PVGRF (BW), SLCMJ PVGRF Asym (%), SLHD L (% leg length), SLHD R (% leg length), SLHD Asym (%), TJ Knee Valgus L, TJ Knee Valgus R, Y-B (% leg length) L, Y-B (% leg length) R, Y-B Asym (%) (BMI = Body mass index; Asym = asymmetry; BW = body weight; SLCMJ = single leg countermovement jump; SLHD = single leg hop for distance; TJ = Tuck Jump; PVGRF = peak vertical ground reaction force; Y-B = y-balance; 75%Hop = 75% horizontal hop and stick; R = right; L = left)	AUC = 0.66, Sensitivity = 0.56, Specificity = 0.74
Jauhiainen et al., 2021	Basketball, floorball	314(57)	Knee, Ankle	LR	Sex, Body mass index, Hamstring flexion non-dominant, KT1000 dominant, Hamstring flexion dominant, Hip flexion peak dominant, Medial knee displacement non-dominant, Vertical ground reaction force non-dominant, Medial knee displacement dominant, Height Knee flexion IC non-dominant, Ankle flexion IC dominant, Leg press one repetition maximum Hip flexion moment non-dominant, Non-dominant knee previous injuries, Knee valgus IC dominant, Knee valgus peak non-dominant, Ankle flexion IC non-dominant, Knee flexion moment peak non-dominant, Vertical ground reaction force dominant, Ankle dorsiflexion dominant, Non-dominant ankle previous injuries, Knee flexion moment peak dominant, Knee valgus IC non-dominant, Generalized joint laxity, Age, Genu Recurvatum Dominant, Dominant ankle previous injuries, Isokinetic flexion non-dominant, Ankle dorsiflexion non-dominant, Anteversion non-dominant, Hip flexion IC non-dominant, Hip flexion IC dominant, Knee flexion peak non-dominant, Knee flexion IC dominant, Hip flexion moment dominant, Knee valgus peak dominant, Navicular drop dominant, Isokinetic flexion dominant, KT1000 non-dominant, Genu Recurvatum Non-Dominant, Exposure, Dominant knee previous injuries, Knee abduction moment peak non-dominant, Knee abduction moment peak dominant, Knee Flexion Peak Dominant, Anteversion dominant, Hip flexion peak non-dominant, Isokinetic extension non-dominant, Isokinetic extension dominant, Navicular drop non-dominant, Hip strength non-dominant, Hip strength dominant, Weigh	AUC = 0.65
Ruiz-Perez et al., 2021	Futsal	139(25)	Lower Extremity	SVM	Player position, Current level of play, Dominant leg, Sex, Age, Body mass (kg), Stature (cm), History of lower extremity soft tissue injury last season, Sleep quality, Physical/emotional exhaustion, Reduced sense of accomplishment, Sport devaluation, Stress control, Influence of sport evaluation, Mental skills, Motivation, Team cohesion, PTISOM-HipAbd-Normalized, PTISOM-HipAdd- Normalized, UnRatio-ISOM-HipAbd/HipAdd, BilaRatio-PTISOM-HipAbd, BilaRatio-PTISOM-HipAdd, Y-Balance-Anterior, Y-Balance-PosteroMedial, Y-Balance-PosteroLateral, BilaRatio-Y-Balance-Anterior, BilaRatio-Y-Balance-PosteroMedial, BilaRatio-Y-Balance-PosteroLateral, Y-Balance-Composite, ROM-HFKF, ROM-HFKE, ROM-HE, ROM-HABD, ROM-HIR, ROM-HER, ROM-KF, ROM-AKDFKE, ROM- AKDFKF, ROM-BIL- HFKF, ROM-BIL- HFKE, ROM-BIL- HE, ROM-BIL- HABD, ROM-BIL-HIR, ROM-BIL- HER, ROM-BIL- KF, ROM-BIL- AKDFKE, ROM-BIL- AKDFKF (Bila: bilateral; Uni: unilateral; ISOM: isometric; PT: peak torque; Abd: abduction; Add: adduction; ROM: range of motion; HFKF: hip flexion with the knee flexed; HFKE: hip flexion with the knee extended; HE: Hip extension; HABD: hip abduction at 90º of hip flexion; HIR: hip internal rotation; HER: hip external rotation; KF: knee flexion; AKDFKE: ankle dorsi-flexion with the knee extended; AKDFKF: ankle dorsi-flexion with the knee flexed; BIL: bilateral ratio)	AUC = 0.77, Sensitivity = 0.66, Specificity = 0.62, F1 = 0.47

Supplementary Table 2. Continue…

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Bogaert et al., 2022	Multiple	161(41)	Lower Extremity	SVM	Gender, Weight, Height, Previous injuries, Use of insoles, Distance covered during the test, 75 auto-extracted statistical features (Not Reported), Root-mean-square ratio, Step regularity, Stride regularity, Sample entropy, Standard deviation of sample entropy, Standard deviation of impact, Step time, Angle in medial–lateral direction during the first two minutes, Angle in medial–lateral direction during the last two minutes, Angle in anterior–posterior direction during the first two minutes, Angle in anterior–posterior direction during the last two minutes	Male AUC = 0.62 Female AUC = 0.65
Jauhiainen et al., 2022	Handball, Soccer	791(60)	Knee	SVM	Age tested, Body mass, Height, Age started elite play, Number of seasons elite play, Match hours avg week, Previous acl, Family acl history, Current acl prevention, Leg press max, Kt1000 predef pull mm, Kt1000 manual pull mm, Ham mobility degrees , Genu recurvatum degrees , Hip anteversion degrees , Knee valgus static , Pelvis forward tilt static, Pelvis l tilt static, Pelvis l rot static, Pelvis width, Femur length , Tibia length , Femur condyle width , Tibia condyle width , Leg length , Quad best, Ham best, H q relation, Single leg squat hip , Single leg squat knee , Single leg drop jump hip, Single leg drop jump knee, Dropjump bilateral, Anterolateral cm , Mediolateral cm , Posterolateral cm , Hip abduction kg , Navicular drop, Gl index, Jump hip flex ic , Jump hip flex max , Jump hip abd ic , Jump hip abd max , Jump hip rot ic , Jump knee flex ic , Jump knee flex max , Jump knee valgus ic , Jump knee valgus max , Jump knee rot ic , Jump ankle plflex ic , Jump ankle plflex max , Jump ankle inv ic , Jump ankle inv max , Jump ankle rot ic , Jump hip mom flex max , Jump hip mom flex max100 , Jump hip mom abd max , Jump hip mom abd max100, Jump knee mom flex max, Jump knee mom flex max100, Jump knee mom abd max, Jump knee mom abd max100, Jump ankle mom dflex max, Jump ankle mom dflex max100, Jump ankle mom inv max, Jump ankle mom inv max100, Jump grf vert max, Jump grf vert time max, Jump grf med max, Jump grf med time max, Jump grf post max, Jump grf post time max , Jump hip flex time max, Jump knee flex time max , Jump knee valgus time max, Jump ankle plflex time max, Jump hip mom flex time max, Jump hip mom abd time max, Jump knee mom flex time max, Jump knee mom abd time max, Jump impulse, Jump fppa ic, Jump fppa max, Jump com height ic, Jump com height min, Jump com height time min, Jump com vertical speed ic, Jump torso flex ic, Jump height, Cut hip flex ic, Cut hip flex max, Cut hip abd ic, Cut hip abd max, Cut hip rot ic, Cut knee flex ic, Cut knee flex max , Cut knee valgus ic, Cut knee valgus max, Cut knee rot ic, Cut ankle plflex ic, Cut ankle plflex max , Cut ankle inv ic , Cut ankle inv max , Cut ankle rot ic , Cut hip mom flex max , Cut hip mom flex max100 , Cut hip mom abd max , Cut hip mom abd max100 , Cut knee mom flex max , Cut knee mom flex max100 , Cut knee mom abd max , Cut knee mom abd max100 , Cut ankle mom dflex max , Cut ankle mom dflex max100 , Cut ankle mom inv max , Cut ankle mom inv max100 , Cut grf vert max , Cut grf vert time max , Cut grf med max , Cut grf med time max , Cut grf post max , Cut grf post time max , Cut hip flex time max , Cut knee flex time max , Cut knee valgus time max , Cut ankle plflex time max , Cut hip mom flex time max , Cut hip mom abd time max , Cut knee mom flex time max , Cut knee mom abd time max , Cut impulse , Cut torso pelvis flex ic , Cut torso pelvis latflex ic , Cut torso pelvis lrot ic , Cut torso ground flex ic , Cut torso ground latflex ic , Cut torso ground lrot ic , Cut torso lrot speed ic , Cut stance time, Cut cutting angle , Cut approach speed ic , Cut foot rrot ic , Cut toe landing ic , Cut cut width com ic , Cut cut depth com ic , Cut cut width pelvis ic , Cut cut depth pelvis ic , Cut momentarm sagittal 40 , Cut momentarm frontal 40 , Cut simplemom frontal maxmom , Cut momentarm frontal maxmom , Cut grf filt maxmom	AUC = 0.63

Supplementary Table 2. **Continue…**

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Huang et al., 2022	Basketball	16(27)	Lower Extremity	dFusionModel	sRPE, Menses, Fatigue, Sleep Quality, Muscle Soreness, Stress Levels, Desire, Urine Protein, Urobilinogen, Urine pH, Urine Specific Gravity, Urine Blood, Urine Ketones, Squat 1RM, 15 m × 17 Shuttle Run, 5.8 m × 6 Shuttle Run, Maximum Vertical Jump, Injury Severity	Minimal LENCI (Precision = 0.93, Sensitivity = 0.92, F2 = 0.92) Mild LENCI (Precision = 0.90, Sensitivity = 0.90, F2 = 0.90)
Lu et al., 2022	Basketball	2103(736)	Lower Extremity Muscle Strains	XGBoost	Recent groin injury, Recent ankle injury, Recent concussion, Recent hamstring injury, Recent back injury, Age, Recent quad injury, Previous injury count, Position, Games played, Games started, Minutes per game, Field goals made per game, Field goal attempts per game, Field goal percent-age, 3-point shots made per game, 3-point shots attempted per game, 3-point percentage, 2- point shots made per game, 2-point shots attempted per game, 2-point percentage, Effective field goal per-centage, Free throws made per game, Free throws attempted per game, Free throw percentage, Offen-sive rebounds per game, Defensive rebounds per game, Total rebounds per game, assists per game, steals per game, Blocks per game, Turnovers per game, Personal fouls per game, Points per game, Player efficiency rating, True shooting percentage, 3-point attempt rate, Free throw attempt rate, Offensive rebound percentage, Defensive rebound percentage, Total rebound percentage, Assist percentage, Steals percentage, Blocks percentage, Turnover percentage, Usage percentage, Offensive win share, Defensive win share, Win shares, Win shares per 48 min, Offensive box ±, Defensive box ±, Box ±, Value over replacement player	AUC = 0.84
Huang et al.,2023	Basketball	17(18)	Lower Extremity	Cost-Sensitive Neural Network (Cost-NN)	Agility: 1-minute double under, 1-minute double under IHR (1min), 1-minute double under HRR (1min), hexagon agility test, hexagon agility test IHR (1min), hexagon agility test, HRR (1min), RPE, urinary protein, urobilinogen, urinary-PH, urinary specific gravity, urinary blood, urinary ke-tones. Speed Endurance: 15 m × 13 shuttle run time, 15 m × 13, shuttle run IHR, 15 m × 13 shuttle run HRR (1min), 15 m × 13 shuttle run BLA (3min), RPE, urinary protein, urobilinogen, urinary-PH, urinary specific gravity, urinary blood, urinary ketones. Strength: 30-sec-ond 35kg squat, 30-second 20kg bench press, 30-second sit up, 30-second back up, CK change, RPE, urinary protein, urobilinogen, urinary-PH, urinary specific gravity, urinary blood, urinary ke-tones. Speed: 3/4 basketball court sprint time, RPE, urinary protein, urobilinogen, urinary-PH, urinary specific gravity, urinary blood, urinary ke-tones. (IHR: instantaneous heart rate; HRR: heart rate recovery; RPE: ratings of perceived exertion; BLA: blood lactate; CK: creatine kinase.)	AUC = 0.86, Precision = 0.64, Sensitivity = 0.87

Supplementary Table 2. Continue…

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Javier Robles-Palazon et al., 2023	Soccer	260(45)	Lower Extremity Soft Tissue	SVM	Player position, Chronological age (y), Age group, Dominant leg, 12 months LE-ST time loss injury history, Years of playing football (y), Training frequency (days), Body mass (kg), Stature (cm), Body mass index (kg/m2), Leg length (cm), Tibia length (cm), Maturity offset, Age at peak height velocity, Anxiety-Trait, Tension, Depression, Anger, Vigour, Fatigue, Confusion, Friendliness, Stress control, Performance evaluation, Motivation, Mental skills, Team cohesion, Global score. Tuck Jump Assessment: FPPA, BIL-FPPA, HF_IC (º), KF_IC (º), AF_IC (º), HF_PF (º), KF_PF (º), AF_PF (º), HF_ROM (º), KF_ROM (º), AF_ROM (º). Drop Vertical Jump: H (cm), CT (ms), RSI (mm/ms), FPPA, BIL FPPA, KMD, BIL-KMD, KASR, KSD (cm), HF_IC (º), KF_IC (º), AF_IC (º), HF_PF (º), KF_PF (º), AF_PF (º), HF_ROM (º), KF_ROM (º), AF_ROM (º). Countermovement Jump: H (cm). Single-leg countermovement jump: H (cm), BIL-H, Take-off pVGRF (N·kg-1), Landing-pVGRF (N·kg-1), pLFT (ms), Take-off BIL-pVGRF, Landing BIL pVGRF, BIL-pLFT, SLJ (cm), SHD (% leg length), SHD-BIL. 10m-Sprint (s), 20m-Sprint (s), 10to20m-Sprint (s), Vmax (m·s-1), M_F0 (N·kg-1), V(0) (m·s-1), Pmax (W·kg-1), DRF (%), FV (N·s·m-1·kg-1), RF-10m (N·kg-1), RFPeak (%),ROM-PHFKF (º), ROM-PHFKE (º), ROM-PHE (º), ROM-PHABD (º), ROM-PHABDHF (º), ROM-PHADD (º), ROM-PHIR (º), ROM-HER (º), ROM-PKF (º), ROM-ADFKE (º), ROM-ADFKF (º), ROM-BIL-PHFKF, ROM-BIL-PHFKE, ROM-BIL-PHE, ROM-BIL-PHABD, ROM-BIL-PHABDHF, ROM-BIL-PHADD, ROM-BIL-PHIR, ROM-BIL-PHER, ROM-BIL-PKF, ROM-BIL-ADFKE, ROM-BIL-ADFKF, YBalance-Anterior (%leg length), YBalance-PosteroMedial (%leg length), YBalance-PosteroLateral (%leg length), BIL-YBalance-Anterior, BIL-YBalance-PosteroMedial, BIL-YBalance-PosteroLateral, YBalance-Composite (%leg length). (SLJ: standing long jump; SHD: single hop for distance; H: height; CT: contact time; RSI: reactive strength index; FPPA: frontal plane projection angle; HF: hip flexio; KF: knee flexion; AF: ankle flexion; IC: initial contact; PF: peak flexion; ROM: range of motion; KSD: knee separation distance; KASR: knee-to-ankle separation ratio; KMD: knee medial displacement; pVGRF: peak vertical ground reaction force; pLFT: peak landing force timing; BIL: bilateral ratio; Vmax: maximal velocity; M_F0: theoretical maximal force; V(0): theoretical maximal velocity; Pmax: maximal power; DRF: decrease inthe ratio of horizontal-to-resultant force; FV: slope of the force-velocity relationship; RF: ratio of the net horizontal-to-resultant force; RFPeak: maximal ratio of horizontal-to-resultant force; ROM: range of motion; PHFKF: passive hip flexion with the knee flexed; PHFKE: passive hip flexion with the knee extended; PHE: passive hip extension; PHABD: passive hip abduction; PHABDHF: passive hip abduction at 90º of hip flexion; PHADD: passive hip adduction; PHIR: passive hip internal rotation; PHER: passive hip external rotation; PKF: passive knee flexion; ADFKE: passive ankle dorsiflexion with the knee extended; ADFKF: passive ankle dorsiflexion with the knee flexed; BIL: bilateral ratio; BIL: bilateral ratio.)	AUC = 0.70, Sensitivity = 0.54, Specificity = 0.74, F1 = 0.38

Supplementary Table 2. Continue…

Reference	Sport	Participants (Injury)	Injury Location	Optimal Model	Candidate Risk Factors	Performance
Kolodziej et al., 2023	Soccer	56(23)	Lower Extremity	LASSO LR	Age, Height, Weight, Postural Control - COP sway (cm), Postural Control - DPSI, Postural Control - Path of platform (mm) , Strength - Trunk (isometric) - Flex (N·m kg−1) , Strength - Trunk (isometric) - Ext (N·m kg−1) , Strength - Trunk (isometric) - Flex + Ext (N·m kg−1) , Strength - Trunk (isometric) - Flex/Ext , Strength - Trunk (isometric) - LatFlex (N·m kg−1) , Strength - Trunk (isometric) - LatFlexr/LatFlexl , Strength - Trunk (isometric) - TransRot (N·m kg−1) , Strength - Trunk (isometric) - TransRotr/TransRotl , Strength - Trunk (isometric) - Core Score (N·m kg−1) , Strength - Hip (isometric) - ABD (N·m kg−1) , Strength - Hip (isometric) - ADD (N·m kg−1) , Strength - Hip (isometric) - ABD/ADD , Knee (isokinetic) - Qcon (N·m kg−1) , Knee (isokinetic) - Qconl/Qconr , Knee (isokinetic) - Hcon (N·m kg−1) , Knee (isokinetic) - Hconl/Hconr , Knee (isokinetic) - Qecc (N·m kg−1) , Knee (isokinetic) - Qeccl/Qeccr , Knee (isokinetic) - Hecc (N·m kg−1) , Knee (isokinetic) - Heccl/Heccr , Knee (isokinetic) - Conventional knee ratio: Hcon/Qcon , Knee (isokinetic) - Functional knee ratio: Hecc/ Qcon , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - Plantarflexion(+)/Dorsalflexion(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - Plantarflexion(+)/Dorsalflexion(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - Eversion(+)/Inversion(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - Eversion(+)/Inversion(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Ankle - External Rotation(+)/Internal Rotation(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - Flexion(+)/Extension(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - Flexion(+)/Extension(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - Adduction(+)/Abduction(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - Adduction(+)/Abduction(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Knee - External Rotation(+)/Internal Rotation(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - Flexion(+)/Extension(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - Flexion(+)/Extension(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - Adduction(+)/Abduction(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - Adduction(+)/Abduction(−) PEAK, Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during SLDL - Hip - External Rotation(+)/Internal Rotation(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Ankle - Plantarflexion(+)/Dorsalflexion(−) IC, Joint kinematics at IC (°) and PEAK (°) during USCC - Ankle - Plantarflexion(+)/Dorsalflexion(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Ankle - Eversion(+)/Inversion(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Ankle - Eversion(+)/Inversion(−) PEAK , Joint kinematics at IC (°) and PEAK	AUC = 0.63, Sensitivity = 0.35, Specificity = 0.79

	(°) during USCC - Ankle - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Ankle - External Rotation(+)/Internal Rotation(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - Flexion(+)/Extension(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - Flexion(+)/Extension(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - Adduction(+)/Abduction(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - Adduction(+)/Abduction(−) PEAK, Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Knee - External Rotation(+)/Internal Rotation(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - Flexion(+)/Extension(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - Flexion(+)/Extension(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - Adduction(+)/Abduction(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - Adduction(+)/Abduction(−) PEAK , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - External Rotation(+)/Internal Rotation(−) IC , Joint kinematics at IC (°) and PEAK (°) during USCC - Hip - External Rotation(+)/Internal Rotation(−) PEAK , PEAK joint moments (Nm/kg) - Ankle - Plantarflexion(+)/Dorsalflexion(−) SLDL , PEAK joint moments (Nm/kg) - Ankle - Plantarflexion(+)/Dorsalflexion(−) USSC , PEAK joint moments (Nm/kg) - Ankle - Eversion(+)/Inversion(−) SLDL , PEAK joint moments (Nm/kg) - Ankle - Eversion(+)/Inversion(−) USSC , PEAK joint moments (Nm/kg) - Ankle - External Rotation(+)/Internal Rotation(−) SLDL , PEAK joint moments (Nm/kg) - Ankle - External Rotation(+)/Internal Rotation(−) USSC, PEAK joint moments (Nm/kg) - Knee - Flexion(+)/Extension(−) SLDL , PEAK joint moments (Nm/kg) - Knee - Flexion(+)/Extension(−) USSC , PEAK joint moments (Nm/kg) - Knee - Adduction(+)/Abduction(−) SLDL , PEAK joint moments (Nm/kg) - Knee - Adduction(+)/Abduction(−) USSC , PEAK joint moments (Nm/kg) - Knee - External Rotation(+)/Internal Rotation(−) SLDL , PEAK joint moments (Nm/kg) - Knee - External Rotation(+)/Internal Rotation(−) USSC , PEAK joint moments (Nm/kg) - Hip - Flexion(+)/Extension(−) SLDL , PEAK joint moments (Nm/kg) - Hip - Flexion(+)/Extension(−) USSC , PEAK joint moments (Nm/kg) - Hip - Adduction(+)/Abduction(−) SLDL , PEAK joint moments (Nm/kg) - Hip - Adduction(+)/Abduction(−) USSC, PEAK joint moments (Nm/kg) - Hip - External Rotation(+)/Internal Rotation(−) SLDL , PEAK joint moments (Nm/kg) - Hip - External Rotation(+)/Internal Rotation(−) USSC , PEAK vGRF (N/kg) - vGRF SLDL , PEAK vGRF (N/kg) - vGRF USSC (ABD, hip abduction; ABD/ADD, ratio between hip abduction and hip adduction; ADD, hip adduction; Conventional knee ratio, ratio between knee flexion concentric and knee extension concentric; Core Score, sum of trunk flexion, trunk extension, trunk lateral flexion right, trunk lateral flexion left, trunk transversal rotation right and trunk transversal rotation left; COP, center of pressure; DPSI, Dynamic Postural Stability Index; Ext, trunk extension; Flex + Ext, sum of trunk flexion and trunk extension; Flex, trunk flexion; Flex/Ext, ratio between trunk flexion and trunk extension; Functional knee ratio, ratio between knee flexion eccentric and

	knee extension concentric; Hcon, knee flexion concentric; Hecc, knee flexion eccentric; LatFlex, trunk lateral flexion; LatFlexl, trunk lateral flexion left; LatFlexr, trunk lateral flexion right; LatFlexr/LatFlexl, ratio between trunk lateral flexion right and trunk lateral flexion left; Qcon, knee extension concentric; Qecc, knee extension eccentric; TransRot, trunk transversal rotation; TransRotl, trunk transversal rotation left; TransRotr, trunk transversal rotation right; TransRotr/TransRotl, ratio between trunk transversal rotation right and trunk transversal rotation left, IC, initial contact: first instance of ground contact phase; kg, kilogram; N, newton; Nm, newton meter; PEAK, peak value: peak value within the first 100 ms after IC; SLDL, single-leg drop landing; USSC, unanticipated side-step cutting; vGRF, vertical ground reaction force)

Supplementary Table 3. Table of machine learning and statistical definitions.

Abbreviation	Full Name	Definition
AUC	Area Under the Receiver Operating Characteristic Curve	AUC quantifies the overall ability of a binary classifier to distinguish between positive and negative classes by computing the area under the ROC curve, with values ranging from 0.5 (random) to 1 (perfect classification).
Precision	Precision	Defined as TP / (TP + FP), precision indicates the proportion of positive identifications that were actually correct. High precision indicates a low false positive rate.
Sensitivity	Sensitivity (Recall, True Positive Rate)	Defined as TP / (TP + FN), it measures the proportion of actual positives correctly identified by the model, reflecting the model's completeness in detecting positives.
Specificity	Specificity (True Negative Rate)	Defined as TN / (TN + FP), it assesses the proportion of actual negatives correctly identified. A higher specificity implies fewer false positives.
DT	Decision Tree	A tree-structured model that splits data based on feature thresholds to predict a target variable. It uses recursive partitioning to maximize information gain or minimize impurity (e.g., Gini or entropy).
RF	Random Forest	An ensemble of decision trees trained on bootstrapped subsets with feature randomness, improving generalization by averaging predictions to reduce overfitting.
SVM	Support vector machine	A supervised classifier that finds the optimal hyperplane to separate classes by maximizing the margin between support vectors, applicable in both linear and non-linear spaces via kernel tricks.
XGBoost	eXtreme Gradient Boosting	An efficient and scalable implementation of gradient boosting that uses second-order derivatives, regularization, and tree pruning for accurate and fast predictive modeling.
ANN	Artificial Neural Networks	A class of models inspired by biological neurons, composed of layers of interconnected nodes (neurons) that learn hierarchical representations through weighted summation and activation functions.
Cost-NN	Cost-Sensitive Neural Network	A neural network trained with misclassification cost weights to penalize minority class errors more heavily, often used in imbalanced data contexts.
dFusionModel	RF-based fusion of XGBoost submodels	A meta-classifier that combines outputs from RF and XGBoost submodels using majority voting or weighted averaging to enhance robustness and accuracy.
LASSO LR	LASSO Logistic Regression	Logistic regression with L1 regularization that shrinks coefficients to zero, performing variable selection and preventing overfitting in high-dimensional settings.