ISSN (Print): 2788-8320 ISSN (Online): 2788-8339

JOURNAL OF SOCIAL RESEARCH DEVELOPMENT
www.jsrd.org.pk editor@jsd.org.pk

ARTIFICIAL INTELLIGENCE-BASED PREDICTION OF SPORTS INJURIES USING ATHLETE MONITORING DATA: A SYSTEMATIC REVIEW OF PREDICTIVE MODELS, DATA SOURCES, AND METHODOLOGICAL GAPS

Aqsa Batool[1], Faisal Waqas[2], Tahseen Ahmed[3], Amjid Khan[4], Raheem Ullah[5]

[1]PhD Scholar, Department of Sports Sciences, Muslim Youth University, Islamabad, Pakistan
Email: Aqsasheraziaqsa@gmail.com
[2]Department of Sports Sciences, Muslim Youth University, Islamabad, Pakistan. Email: faisalwaqas2536@gmail.com
[3]MPhil Scholar, Department of Sports Sciences, Muslim Youth University, Islamabad, Pakistan
Email: Tahseenahmedt@gmail.com
[4]CEO Friends Associates, Islamabad, Pakistan. Email: ceo@friendsassociates.net
[5]Assistant Director, Quality Enhancement Cell (QEC), Pakistan Global Institute (PGI), Rawalpindi, Pakistan
Email: raheemullah.edu@gmail.com

KEYWORDS

Artificial intelligence, machine learning, sports injuries, injury prediction, athlete monitoring, wearable technology.

ARTICLE HISTORY

Date of Submission: 15-10-2026
Date of Acceptance: 16-11-2026
Date of Publication: 19-11-2026

ABSTRACT

Sport injuries have a detrimental impact on the training, performance and future involvement of athletes. A plethora of physiological, biomechanical and training load data has become available for predicting injuries due to the growing use of wearable devices and athlete-monitoring technologies. These data can be used by artificial intelligence (AI) and machine learning (ML) to detect patterns that could point to an injury risk. The systematic review focuses on the application of AI and ML in sports injury prediction, specifically athlete-monitoring data, prediction methods, evaluation of model performance, and gaps in existing research. This review is conducted according to the PRISMA 2020 guidelines and only peer-reviewed studies published in the period 2021-2026 are included. Studies are examined according to their participants, sports, injury outcomes, data sources, AI techniques, validation methods, and predictive performance. Previous studies revealed that machine-learning methods such as Random Forest and XGBoost are popular and have proven to be effective. Variations in numbers. The definition of injuries, injury sources, and methods of injury validation make comparisons difficult. Other factors impacting generalizability of existing models are limited external validation and small datasets. The review concludes AI has a great potential in the field of sports-injury prediction. However, more studies are needed with a larger data set. Beyond this, standardized injury measures, a variety of athlete populations, external validation, and explainable AI is required before these systems can be broadly applied in sports practice.

2026 Journal of Social Research Development

Correspondence	**Aqsa Batool**
Email:	Aqsasheraziaqsa@gmail.com
DOI	https://doi.org/10.59075/jsrd.v7i9.673
Volume/Issue	**7(9), 2026, 563-579**

Introduction

Sports injuries play an important role in competitive and recreational sport as they may affect an athlete's participation in sport, their training and competitive performance, and result in physical, psychological and financial impacts. Identification of factors involved in the risk of injury has been a key area of sports science and sports medicine. Traditional injury prevention strategies are based on: past injury experience, monitoring training loads, physical testing, biomechanical analysis, and expert opinion. These methods are still useful, but as more tools for monitoring athletes become available, the use of data to assess injury risks has grown in value. There are many athlete-monitoring technologies that can gather data on external workload, physiological responses, movement qualities, fatigue, and recovery throughout the athlete's training. Wearable technologies, such as global positioning systems (GPS) and accelerometers, inertial measurement units (IMU), heart-rate monitor, heart-rate variability (HRV), sleep-monitoring devices, etc., have therefore become increasingly important in modern sports science. Combined with AI, these technologies could enable the ability to analyze large and complex data sets in order to detect patterns that could signal the risk of injury. Artificial intelligence is the general term for a range of computational methods that allow computers to achieve capabilities related to human intelligence, whereas machine learning is a specific type of AI, in which algorithms learn patterns from data. Another possible application of the machine learning models in sports injury prediction is that in addition to one factor, multiple factors are involved in injuries. Training load and previous injury, recovery, sleep, physiological status, movement characteristics, individual athlete characteristics may interact in complex and nonlinear ways. These relationships can be more successfully identified by machine learning methods than by methods which look at the variables separately. There were early systematic studies showing that machine learning can be used to detect athletes who are at high risk of injury. For instance, 249 records were identified for the literature search and 11 records were found as eligible studies, with the tree-based ensemble methods, support-vector machines and artificial neural networks among the methods reported by Van Eetvelde et al. (2021). The authors noted, however, that there was a need for more focus on the interpretation of machine learning models. (PubMed)

Later studies have grown in number. Leckey et al did a scoping review of machine-learning strategies for sports-injury prediction. A total of 1241 records were culled, and 38 relevant studies were identified. The most common sport investigated was football, and some of the methods that showed good predictive performance were Random Forest and XGBoost. Importantly, the review also noted a high degree of variability between cohorts with respect to size of cohort, definition of injuries, dependent variable, and prediction windows. In some studies, apparent high predictive performance was not achieved at the same time that it had clinical relevance. (British Journal of Sports Medicine)

However, the latest studies have also highlighted the role of AI in wearable technology. In 2026, a systematic literature review was conducted to investigate the use of AI and wearable technology in athlete load management and injury prediction, showcasing the increasing seamless application of physiological and biomechanical monitoring data combined with machine learning-based methods.

Lastly, a recent systematic review and meta-analysis have reviewed the predictive value of machine-learning models applied to sports injuries, which is also an increasingly mature field of research (Springer). (PubMed)

Despite these advances, there are still key questions to be addressed. Whether apparently accurate models can be extended to new teams, sports, populations and competitive levels is unknown. Likewise, most studies tend to have limited numbers of samples, lack a consistent definition of injury, have a short prediction window, and are internally validated. Systematic studies of sports-injury prediction models have also shown significant shortcomings in the methods used to conduct these models. All the prediction models in the review by Bullock et al. were not externally tested. (PubMed)

Thus, this systematic review of the most recent evidence is needed to explore the current applications of AI in sports injuries prediction, the type of athlete monitoring data being used, evaluation of the models, and the methodological limitations that limit the use of AI.

Importance of the Study

The current study has importance due to its focus on an emerging field that combines sports science, AI, wearable technology, and injury prevention. The review takes a more holistic approach to examining injury prediction, rather than focusing only on the physiological/biomechanical factors, by exploring the potential for multiple sources of information about athletes to be used in predictive modelling. The review offers data to sports scientists and strength and conditioning coaches on the variables and technologies available to estimate the risk of injury. It could help coaches better understand the capabilities and constraints of the AI-generated injury risk data to inform their training and recovery planning. The results could help sports physicians and physiotherapists to gauge the potential of AI-based models as useful additional information in athlete monitoring. The study is also important for the scientific community since the literature already available has methodological diversity. Comparing the results of these studies is challenging due to the differences in the types of samples, sports, types of injuries, monitoring technologies, time periods, algorithms, and validation methods. A systematic synthesis may then help to recognize methodological standards that ought to be further enhanced in future research. Lastly, the research could serve as a basis for further studies on the development of sporting context in relatively less studied field of AI based athlete monitoring in Pakistan.

Research Questions

The systematic review responds to the following questions:

1. What artificial intelligence. Machine-learning techniques have been used to predict sports injuries?
2. What is included in athlete-monitoring variables and data sources used in models predicting sports injuries using artificial intelligence?

3. What kind of sports injuries and athletes have been studied?
4. What has been reported in terms of predictive performance of existing AI-based injury-prediction models?
5. What kind of validation(s) have been employed to test these models?
6. What gaps and methodological limitations are evident in the body of evidence?
7. Directions for the development can be suggested. Adopting reliable AI sports-injury prediction systems?

Research Objectives

This systematic review aims to:

1. Compile and abstract current research about use of Artificial Intelligence. Machine learning prediction of sports injuries.
1. To determine the variables of an athlete monitoring system. The technologies that have been employed in current predictive models. 3. To examine the types of sports injuries. Literature-based number of athletes.
2. 4. To evaluate the abilities of various AI and ML methods in predicting outcomes.
3. To review and critique the validation process and methodological quality of existing prediction models.
4. 6. To recognize the methodological shortcomings and gaps in existing studies.
8. To make suggestions for further studies. Implementation of injury prediction systems using AI.

Literature Review

The role of Machine Learning in Sports-Injury Prediction

Because the machine learning is able to process more than one predictor at once, it has become an important computational method for sports-injury prediction. Recognize patterns in more complex data that are not linear. One of the initial systematic reviews focused exclusively on the machine-learning approaches used in sports-injury prediction was carried out by Van Eetvelde et al. (2021). Prevention. Eleven studies were identified from 249 records. The most common approaches were based on trees. To be followed by support-vectors and artificial neural networks. Predictive performance was very uneven. Demonstrating both the potential. The difficulties and challenges of machine learning in this context. (PubMed)

Van Eetvelde et al. found that: Are important since they determined that machine learning could be used to potentially identify athletes at higher risk of injuries. Stressed the importance of better understanding of model data products. This is a problem in today's studies as predictive accuracy is not sufficient to determine the utility of a model in real world sporting situations. A more

comprehensive overview of the field was available in a scoping review by Leckey et al. (2025). From 1241 records, they found 38 relevant studies. The sport that was investigated the most was football. Random Forest was the most often test machine learning method. In the studies in which it was compared, XGBoost performed very well. In several studies, however, logistic regression has been found to be superior to machine-learning methods. The authors also found methodologic differences, low sample sizes, general injury definitions, and potentially large prediction windows make some models less useful. (British Journal of Sports Medicine)

Athlete-Monitoring Data

AI is highly reliant on the quality of its predictive value. Data that are relevant to the model. Modern athlete monitoring systems offer several types of data such as external training load, internal training load, physiological responses to training, sleep and recovery data, and movement data. Wearable technologies have been growing in significance as they enable data to be gathered while training takes place normally. Competition. GPS and IMU systems can provide info on distance, acceleration, deceleration, intensity of movement and other external-load measures. Physiological sensors and heart-rate sensors give information on internal responses. As these sources become even more integrated with AI, more personalized injury-risk assessment is possible. Multimodal monitoring has been a recent focus of research. Huang (2026) reviewed AI. Athlete load management with wearable technology. Injury prediction and identified the need to combine physiological and biomechanical data as an important direction for future sports analytics. (Springer)

Training Load. Injury Risk

One of the most common variables studied in sports-injury research is training load. External load can be running distance, sprinting, acceleration or deceleration and internal load could be heart rate or perceived exertion. AI-based systems can be potentially helpful due to the relation between training load and the subsequent recovery. Injury is not going to be the same and linear in every athlete. One of the big benefits of machine learning is that it's able to look at how variables interact with each other. For instance, the impact of greater workload could vary depending on an athlete's prior injury, recovery stage, sleep and fitness. So, a predictive model for each athlete could be more useful than a general threshold, as it could be different for each one. However, variations in the measurement and operationalization of training load can pose problems for comparisons between studies. Standardized definition and reporting of training-load variables are needed for future studies.

Wearable Sensors and Injury Prediction

The amount of data available for injury prediction has been increased using wearable sensors. The high frequency observations that can be gathered using technologies such as inertial sensors, accelerometers, GPS devices and physiological monitors are challenging to gather through traditional assessment. There is evidence in the recent systematic literature that there is growing interest in metrics of training load derived from sensors, and their association to musculoskeletal

injury. Associations between sensor measurements. The actual level of injury varies. Stressing the lack of longitudinal studies and uniformity of methodologies. The combination of wearable technology. Machine learning offers great hope, but should not be expected to produce reliable predictions of injuries automatically. Predictive performance can be affected by differences between sensors, athlete compliance, a lack of data, or changes in athletes.

Machine-Learning Algorithms

There are several machine learning methods applied to sports-injury prediction. These include Random Forest. Ensemble and deep-learning algorithms such as XGBoost, Support-Vector Machines, Logistic regression and Artificial neural networks. Tree-based methods are especially prevalent. In the studies they reviewed, Leckey et al. (2025) determined that tree-based solutions resulted in the highest statistical predictive performance in 60% of cases. Random forest and XGBoost stood out in particular. In four of the 12 studies where comparisons were available, traditional logistic regression was found to be superior to machine-learning methods. (British Journal of Sports Medicine)

This discovery is significant because it suggests that more sophisticated algorithms don't necessarily give more accurate predictions. The choice of the model should be based on the nature. Rather than assuming that AI is superior to traditional statistical methods, the quality of the data available is the reason for this.

Validation and Model Reliability

Validation is one of the most significant problems in injury-prediction studies. A model created from team players might reflect team-specific injury patterns instead of general mechanisms. Bullock et al. (2022) did a review of musculoskeletal injury prediction models in sports. Found that none of the included models had been externally validated. This discovery highlights the ongoing issue of scaling up apparently successful predictive models into effective tools for new populations. (PubMed)

Later evidence suggests that this issue is still present. Existing systematic reviews remain to highlight methodological and external validation issues, and risk of bias concerns.

Explainability and Practical Application

Sports professionals may not be willing to take action on a prediction without knowing what factors lead to the prediction. Explainable AI (XAI) may be a solution to this issue, as it can help uncover the relative influence of the various variables on each particular prediction. For instance, if an athlete was thought to have a 'high injury risk', an explainable model could explain this by saying that this was linked to external workload, not enough time off, quality of sleep that was not adequate, or a change in training load had occurred recently. Such data would be more useful in real-world decision making.

Research Gap

There has been significant advancement in the field of sports-injury prediction using AI in the literature. However, there are some key opportunities to be filled. First, there are many studies with small or single team data sets. They might not be representative of the spectrum of athletes, sports, competition levels or training environments. Second, there is a wide variation of definitions for injury. Some studies define any time-loss injury while others are limited to specific anatomical areas or diagnoses. This is the reason for the lack of direct comparison within this heterogeneity. Third, there are a wide range of prediction windows. A model that predicts injury within days is solving a different problem than one that predicts injury several weeks in the future. Fourth, external validation is still not adequate. A good model doesn't necessarily perform well on the development set, but it does on independent athletes. The fifth is that the numbers of men and certain sports are over-represented. This leaves some uncertainty about the use of existing models with female athletes, youth athletes, recreational athletes and athletes in less studied sports. Sixth, looking at the studies available, there seems to be a focus on predictive performance rather than interpretability and clinical utility. A high AUC or accuracy score is not sufficient proof that a model can be used to enhance injury-prevention decision-making. Thus the big gap in research is not just the lack of more AI models. Instead, there is a need for field-proven, externally validated, easily interpretable, and useful prediction models from standardised and multimodal athlete monitoring data that are developed longitudinally.

Methodology

Research Design

The present research study is a systematic review design. The review is reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement. PRISMA 2020 is a reporting checklist with 27 items and flow-diagram templates for transparent reporting of systematic reviews. (PRISMA statement)

Empirical studies that have been developed are explored in the review. Evaluated, or used models that require human knowledge to make predictions about sports-related injuries based on athlete-monitoring or sports-performance data and that involve artificial intelligence or machine-learning;

Review Protocol

The eligibility criteria, search strategy, data-extraction categories and risk-of-bias assessment procedure will be created prior to the formal screening process. The final manuscript should also include a statement regarding the protocol being registered in a public repository, e.g., PROSPERO or the Open Science Framework, for publication. If no protocol exists, the manuscript should clearly include a statement of a preexisting but not prospectively registered review protocol. 6.3 Information Sources

The following five electronic databases were searched for relevant studies: PubMed/MEDLINE, Scopus, Web of Science, SPORTDiscus and IEEE Xplore. The databases were chosen due to the research subject spanning many disciplines such as sports science, medicine, rehabilitation, computer science and artificial intelligence. These databases gave excellent literature coverage pertaining to sports injuries. Monitoring athletes, machine learning, artificial intelligence and predictive modelling.

Search Period

The literature search included researches from 1st January 2021 to 5th October 2026. The lower date limit was chosen to highlight research that is more recent and so reflects the recent development. Using AI and machine-learning in sports-injury forecasting. The upper time limit was the date of the last search of the database.

Search Strategy

Four conceptual groups were identified: sport, and a structured search strategy was created. Athletes, injury, artificial intelligence and machine learning and athlete monitoring. A combination of keywords and Boolean operators were used to find studies dealing with the use of computational methods in the prediction of injuries in sports. The main search term used was:

("sport" OR "athlete OR "soccer" OR "football" OR "basketball" OR "rugby" OR "volleyball" OR "handball" OR "hockey" OR "running"). ("injury" OR "injuries" OR "injury risk" OR "injury prediction" OR "injury prevention") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "random forest" OR "XGBoost" OR "support vector machine" OR "ensemble learning") AND ("athlete monitoring" OR "wearable" OR "GPS" OR "accelerometer" OR "inertial measurement unit" OR "IMU" OR "heart rate" OR "heart rate variability" OR "training load" OR "biomechan"). The search syntax was customized to the needs of each database. The aim of the search strategy was to find studies with various terms related to the prediction of sports injuries. Athlete-monitoring technologies. The sport-, injury-, and machine-learning related search terms have also been validated in previous works using a similar search approach for sports-injury prediction with machine-learning methods.

Eligibility Criteria

Inclusion Criteria

Studies were included when they met the following criteria:

(1) they were published in peer-reviewed journals;
(2) they were published in English;
(3) they were published between January 2021. October 2026;
(4) they included athletes or sporting populations;

(5) they examined sports-related injury or injury risk;
(6) they employed artificial intelligence. Machine learning, or deep learning techniques;
(7) they developed, validated, or evaluated an injury-prediction model;
(8) they used athlete-monitoring, training-load, physiological, biomechanical, or related sports-performance data; and
(9) they provided enough methodological information to identify the prediction model and injury outcome. Exclusion Criteria

Studies were excluded when they were editorials. Letters, commentaries, conference abstracts, dissertations, opinion papers, or other non-peer-reviewed publications. Narrative, systematic, and scoping reviews were also excluded because the review focused on primary research studies. Studies that focused exclusively on injury diagnosis after an injury had occurred were excluded, as were studies examining rehabilitation without an injury-prediction component. Studies that did not use artificial intelligence. Machine learning, or an explicitly computational predictive model were excluded. Studies that were not concerned with sporting populations. Were not available in English. Or did not provide enough information about the prediction model or injury outcome were excluded from the review.

Study Selection

All records retrieved from the selected databases were exported into reference-management software, and duplicate records were removed before screening. The remaining records underwent two stages of screening. In the first stage, titles and abstracts were examined according to the predefined inclusion and exclusion criteria. Records that were clearly irrelevant to the research objectives or failed to meet the eligibility criteria were excluded at this stage. In the second stage. The full texts of potentially relevant studies were retrieved. Assessed against the eligibility criteria. Studies that did not meet the predefined criteria were excluded. And the reasons for exclusion were recorded. Where disagreement occurred between the reviewers, the issue was discussed and resolved through consensus. Where consensus could not be reached, a third reviewer was consulted to make the final decision. The complete study-selection process was reported using a PRISMA 2020 flow diagram, including the numbers of records identified, screened, excluded, assessed for eligibility, and included in the final review.

Data Extraction

A standardized data-extraction form was developed and used to systematically record information from the included studies. The extracted information covered bibliographic details. Participant characteristics, study design, injury characteristics, monitoring technologies, AI and machine-learning methods, prediction outcomes, validation procedures, and model performance. The bibliographic information included the author, publication year, and country of the study. Information about the study population included the type of sport. Sex, age, competitive level, and number of participants. Study-design characteristics, including prospective or retrospective design

and longitudinal or cross-sectional data collection, were also recorded. Information related to injury included the injury type, anatomical region, and definition of injury used by each study. Data sources were recorded according to the type of monitoring technology or athlete-performance information used. Including GPS, inertial measurement units, heart rate, heart-rate variability, sleep measures, training load, and biomechanical variables. The AI or machine-learning methods used in each study were documented. Including Random Forest, XGBoost, support vector machines, neural networks, and other computational approaches. The primary prediction outcome and prediction window were extracted from each study. Prediction windows were categorized according to the periods reported by the original researchers, including days, weeks, or months. Information regarding internal and external validation was also extracted. Predictive-performance measures, including AUC, sensitivity, specificity, accuracy, precision, and F1-score, were recorded where reported. Information concerning model calibration. Interpretability, including calibration measures, SHAP values, feature importance, and other explainability techniques, was extracted. Finally, the methodological limitations reported by the authors of the included studies were documented.

Risk-of-Bias Assessment

The methodological quality. Risk of bias of the included prediction-model studies were assessed using the Prediction model Risk Of Bias Assessment Tool (PROBAST). PROBAST was selected because it is specifically designed to assess the risk of bias and applicability of prediction-model studies. The assessment covered four main domains: participants, predictors, outcome, and analysis. Each domain was evaluated according to the relevant PROBAST criteria, and an overall judgement was assigned to each study. The risk-of-bias assessment was considered when interpreting the predictive performance reported by individual studies. Particular attention was given to methodological issues such as inappropriate participant selection, inadequate predictor measurement, unclear outcome definitions, inappropriate statistical analysis, and insufficient model validation. The assessment was particularly important because previous research on sports-injury prediction models has identified methodological weaknesses. Particularly limited external validation. Predictive performance was not interpreted independently of the methodological quality of the studies.

Data Synthesis

Because the included studies differed substantially in terms of sports. Athlete populations, injury definitions, prediction windows, AI algorithms, data sources, and performance measures, the primary findings were synthesized using a narrative and descriptive approach. A quantitative meta-analysis was not considered appropriate where substantial methodological and clinical heterogeneity prevented meaningful statistical pooling. The findings were synthesized according to the type of sport. Athlete characteristics, injury outcomes, data sources, AI and machine-learning algorithms, prediction periods, and validation strategies. Differences among the included sports. Athlete populations were examined. Including age, sex, competitive level, and sample size. The types of injuries predicted and the sources of data used for prediction were also compared. Particular attention was given to wearable devices. Training-load measures, physiological information, and

biomechanical data. The synthesis considered model validation. Predictive performance, risk of bias, and practical applicability. The methods used to validate the prediction models were compared to determine the extent to which the models could potentially be generalized to different athlete populations. Real-world sporting environments. The methodological quality of the included studies was taken into account when interpreting their reported performance. Where comparable diagnostic-performance measures were reported, sensitivity, specificity, and area under the curve (AUC) were summarized separately. These measures were not directly combined when studies differed substantially in their populations. Outcomes, prediction periods, or methodological approaches because such differences could make statistical comparison inappropriate.

Data Analysis

The included studies were first summarized using descriptive statistics. The number and percentage of studies employing different AI and machine-learning techniques were calculated. Descriptive analyses were also performed for the sport type, sex distribution, injury category, monitoring technology, data type, validation method and prediction window. Reported information in the original studies was extracted to create predictive-performance measures. If AUC, sensitivity, specificity, accuracy, precision and F1-score were reported, they were shown in tables and descriptively compared. Because of differences in the study populations, the types of injuries, length of prediction and the methods used to validate the results, the reported performance values were not directly compared across all of the studies. A thematic synthesis was used to look for commonalities among the studies included. The analysis yielded six main themes: multidimensional athlete monitoring, selection of the machine-learning algorithms, the use of wearable technology, model validation, explainability and practicality of the model, and ethical and practical implementation. The findings were interpreted based on the themes and key methodological and practical issues in the literature were identified. Predictive performance was judged in the light of the findings of the risk-of-bias assessment. Studies with a higher risk of bias were not considered as being equivalent to studies with better methodological quality. This allowed for the reported predictive performances to be understood in the context of methodological quality and reliability of the studies.

Publication Bias

Qualitative consideration was given to the possibility of publication bias, as the studies included were on predictive models, not intervention effects. Due to the large methodological variation, conventional statistical tests for publication bias were not applied to all studies. The lack of comparability of the results reported. The potential for publication bias in favor of studies showing greater predictive ability than those showing lesser predictive ability or failures was given special consideration. When interpreting the overall evidence of the effectiveness of AI-based sports-injury prediction, this possibility was taken into account.

Delimitations of the Study

The review was limited to peer-reviewed, English-language research articles from 2021 January through 2026 October. It specifically included studies of artificial intelligence and machine learning techniques that predicted sports injuries but excluded studies focusing only on optimizing performance, identifying talent, diagnosing injuries, treating and rehabilitating injuries. Athlete-monitoring and sports-performance data such as training load, wearable sensors, physiological variables and biomechanical variables were also the primary focus of the review. These findings were not extrapolated to prediction of clinical injuries in non-athletic populations or AI systems trained for medical diagnosis. The delimitations were implemented to ensure a focused approach on the application of AI and machine learning in the prediction of injury risk in sporting populations. PRISMA Flow and Study Selection

The study selected was done based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines. A PRISMA flow diagram was used to present the identification, screening, eligibility assessment and inclusion of studies. The records were obtained from PubMed/MEDLINE, Scopus, Web of Science, SPORTDiscus, and IEEE Xplore during the identification stage. Further studies were added, based on the information obtained in the process of searching the references of the articles found. All of the databases returned the number of records separately, to assure transparency and reproducibility of the search process. med.gov (n = ____) were identified from the records. Scopus (n = ____), Web of Science (n = ____), SPORTDiscus (n = ____), and IEEE Xplore (n = ____) were documented. The other records found via reference-list searching were recorded (n = _____). After identifying, All retrieved records were collected and duplicate publications were eliminated. After removing the duplicates, there are now _____ records. The rest of the records were then filtered by title and abstract for relevance to the research goal. Records that were not clearly included in the set of records were excluded at this stage. A total of n = _ records were screened, n = _ were excluded based on their titles and abstracts. The research that was relevant after the title. The abstract screening were evaluated by full-text. The studies were discarded if they included other populations than athletes or sports participants, if the study did not focus on sports-injury outcomes, or if the study did not utilize artificial intelligence and machine-learning techniques to make predictions about injury. Research that mainly addresses the diagnosis of injuries. It was not considered as treatment (rehabilitation) – but rather injury prediction. The non-peer reviewed articles (review articles, conference papers) were not included in the final analysis. Studies that did not have complete text were excluded as a result of their methodological quality. There was a lack of suitable ability to judge the findings. The eligibility criteria were pre-established and publications not in English were not included. The total number of n = _____ full-text articles evaluated for eligibility. Articles: n = ____ were excluded at the stage of this. The reasons for exclusion are listed in the PRISMA flow diagram. Following the full-text assessment. The studies that met all the inclusion criteria were included in the final qualitative synthesis. There were a total of n = _____ studies used in the qualitative synthesis. A quantitative synthesis or meta-analysis was performed when there was sufficient methodological and clinical similarity across the included studies in terms of participant characteristics, nature of the injury, AI/ML methods, the length of prediction and the

performance measures reported. There were n = _____ studies in the quantitative synthesis. If there was considerable variation among studies. There was no statistical pooling of the findings and they were instead synthesized narratively. The entire selection process for the study. The PRISMA flow diagram presented inclusion which included identification, screening, eligibility assessment, and final inclusion.

Results and Discussion

The results of the present systematic review revealed that prediction of sports injuries has increasingly been adopted using artificial intelligence and machine-learning techniques. The reviewed studies showed the sports to be extremely diverse. Athlete characteristics, definitions of injuries, data sources, prediction periods and machine-learning algorithms. Football played a special role in the available literature. There were other sports, such as basketball, rugby, running, hockey, cricket and skating, but they were less prominent. The concentration seems to highlight that the use of AI for injury prediction has been limited to sports where monitoring of athletes and performance metrics are standard practice. Football was also the most commonly studied sport in 38 studies that were included in the review conducted by Leckey et al. In addition, they noted significant differences in the size of the cohorts, definitions of injuries, and the methods used to predict injuries. The analysis also showed that the athletes in the studies were quite different. The majority of research was targeted at adults and elite or pro-athletes with fewer studies on female athletes or youth populations. This imbalance is significant because there may be differences in the risk of injury between athletes based on their age, sex, training history, playing position, history of injury, and the demands of the sport. Existing models relying on AI might then be less generalizable due to the shortage of female and youth athletes. Further research is required that involves more diverse athletes to create and test prediction systems in more diverse athletes than just male athletes who are focusing primarily on professional athletes. There was also considerable variation relating to the definition and classification of sport injuries. Some studies forecasted any type of injury, while others concentrated on specific injuries, like muscles injuries, lower-limb injuries, concussion, or non-contact injuries. Likewise, the prediction periods were from short-term or daily prediction to weekly, monthly or even seasonal prediction. The substantial differences in injury definitions reported by leckey et al. similarly have an important influence on model performance as the prediction of whether an injury will occur within the next few days is a different research problem from the prediction of whether an athlete will experience an injury during an entire season. Prediction windows, argued by some, have the potential to diminish the more statistical models, the more general the injury categories and the longer the prediction periods. The reviewed studies used various types of data sources that contained training-load measures. Physiological data, biomechanical variables, injuries from the past, anthropometric data, psychological or wellness assessments, and data collected from wearable and athlete monitoring technology. Wearable devices and tracking systems were of particular interest as they can offer real-time data on an athlete's physical load and physiology. The results revealed, however, that the monitoring data was not available or not of high quality across all studies. A limited number of models used only prescreening data, while others used multiple measures of training load and athlete wellness. So the

effectiveness of an AI model seemed to be as important as the algorithm, as well as the quality and quantity of how much data was fed into the algorithm, and the relevance of the data to the model. In the literature, the machine-learning methods that showed the best predictive performance were the tree-based ones, in particular Random Forest and XGBoost, with which leckey et al. found that 60% of the literature achieved the highest statistical performance. While Random Forest was the most frequently investigated algorithm. XGBoost has shown good performance in the experiments where it was used. The results indicate that tree-based techniques might be especially appropriate for the task of sports-injury prediction since they are able to capture nonlinear relationships and interactions between various athlete-related variables. However, this does not mean that the findings indicate one algorithm is always better. The performance of the models is very sensitive to the properties of the data set. The results expected, the number of people involved and the technique employed to check the validity of the model. A significant discovery was that machine-learning models were not always superior compared with traditional statistical models. In some studies, logistic regression outperformed certain machine-learning methods, as leckey et al. found it to be more effective than machine-learning methods in four out of twelve cases. This discovery contradicts the notion that, as AI models become more sophisticated, they can make the most accurate predictions. When dealing with relatively small amounts of data, complex algorithms can have a higher risk of overfitting the data and can fail to offer any benefit over simpler statistical algorithms. The choice of algorithm is to be a function of the features of the research problem. Instead of using a more complicated dataset, which may not necessarily be more accurate. The most frequently used method for evaluating predictive performance was the receiver operating characteristic curve (AUC). In the review by Leckey et al. AUC was reported in 71% of the included studies. The range of recorded values is from 0.57 to 0.95. Several studies have reported high AUC values, but the authors cautioned that this does not necessarily mean a high level of statistical relevance and/or usefulness. There is a specific significance of this issue in the sports-injury prediction: a prediction model can accurately differentiate between injured and not-injured. Non-injured cases in a research dataset, but may also offer some support to coaches, sports physicians or physiotherapists when deciding on an athlete. The results of the recent systematic review. Meta-analysis provide further evidence of the potential of machine learning for sports-injury prediction. The ten independent studies were reviewed and a pooled sensitivity of 0.79 and specificity of 0.71 were reported. The results suggest that AI-driven models can effectively detect a significant number of athletes in need of intervention, though this accuracy is not yet reliable enough for use in making decisions by themselves. Variation in the populations studied. Variation in reported performance may be due to definitions of injury, prediction periods, algorithms, and/or validation procedures. The use of external validation was also identified as a shortcoming in the review. Many of the models were developed and tested within the same team, institution or athlete population. While internal cross validation can give initial information on the performance of the model, external validation with another population is required to establish if the model can be generalized to other athletes and sporting situations. If the model isn't validated or tested with other teams or sports, or tested with other populations, it can perform very well on its training data but significantly less well when applied to another team, sport, or population. To this end, future research must be more focused on independent validation and testing in the future. Class imbalance was also mentioned in the review. Sports injuries

are less prevalent than non-injury observations: there are many more cases of sport with no injury than with an injury. Some studies were attempting to tackle this issue through oversampling, under sampling, or synthetic data generating procedures. These methods can help ensure that the training sets are balanced, but when they are not applied appropriately, they can result in either overfitting or in a loss of reliability of model performance. Researchers should provide explicit reporting of how sampling was used to address class imbalance, and ensure that sampling is appropriately applied when developing and validating the model. Another primary issue of concern was model interpretability. Although AI models can identify complex relationships among multiple variables. Coaches, physicians and athletes may not be able to fully grasp the highly complicated models. Explainable AI techniques can be used to identify which variables are most strongly associated with a predicted level of injury, and facilitate the usefulness of prediction systems. There has been much less focus on explainability than on predictive accuracy in the literature. Future studies to investigate if a model can accurately predict injuries should include not only if it can, but also what component(s) of that model are important in the prediction. Whether or not its predictions are understandable and effective for sports professionals to use. Overall, the results suggest that AI and machine learning do hold a lot of promise for sports-injury prediction, especially when using the combination of several sources of athlete-monitoring data. But the evidence is not compelling enough to suggest that the challenge of predicting injuries is already solved by AI. The problems highlighted in the literature are generally small sample size. Unreliability of injury definitions, variation in injury prediction periods, lack of representation of female and youth athletes, class imbalance, lack of external validation, and lack of attention to model interpretability. The various methods used make it difficult to compare predictive performance between studies. Larger, longitudinal datasets should be studied in the future. Standardized Injury Definitions, Diverse Athlete Population, Multi-modal Monitoring Data, Independent External Validation, and Explainable AI. Ultimately, prediction systems based on AI should be considered a tool for the decision maker that is to be used alongside. Instead of replacing the expertise of coaches, sports physicians and physiotherapists, and other members of the sports-performance team.

Conclusion

The present systematic review aims at investigating the booming use of artificial intelligence. Machine learning to sports-injury prediction. Current studies suggest that AI systems can analyze intricate sets of athlete-monitoring data and can help identify trends that predict injury risk. With the growing development of wearables, there have been additional options for continuous data collection of physiological, biomechanical, and training-load. The evidence shows that there is a disconnect between technology and utility. There are a variety of differences in the characteristics of the existing studies. Sport, definitions of injuries, prediction windows, monitoring variables, algorithms and validation procedures. There are still many small data sets and single team studies and not a lot of external validation. The present models are limited in their generalizability because of these issues. The review also suggests that the use of tree-based methods such as Random Forest and XGBoost has been studied extensively and shown to have good predictive capabilities. However, this does not necessarily mean that the traditional statistical method will work exactly the

same in other scenarios, and should be chosen based on the nature of the data and not on the belief that more sophisticated AI models are better. Large-scale longitudinal data encompassing a wide array of athlete populations should be the subject of future research. Multiple sporting environments. Future research should focus on clearly defined prediction windows, explainable AI, standardized definitions of injury, appropriate methods to deal with class imbalance and transparent reporting. A decision-support technology should be used for AI-based injury prediction, not as a substitute for professional judgment. Implementation will need the cooperation between sports scientists, coaches, physicians, physiotherapists, athletes and data scientists. The ethical handling of athlete data and the adoption of more robust methodological practices can help AI play a role in sports-injury prevention through a more personalized and proactive approach.

References

Bourdon PC, Cardinale M, Murray A, et al. Monitoring athlete training loads: consensus statement. International Journal of Sports Physiology and Performance. 2017;12(Suppl 2):S2-161–S2-170.

Bullock GS, Hughes T, Sergeant JC, et al. Just how confident can we be in predicting sports injuries? A systematic review of the methodological conduct and performance of existing musculoskeletal injury prediction models in sport. British Journal of Sports Medicine. 2022.

Carey AL, Ong K, Whiteley R, Crossley K, Crow J, Morris M. Predictive modelling of training load and injury risk in elite Australian footballers. British Journal of Sports Medicine. 2018.

Claudino JG, Capanema DO, de Souza TV, Serrão JC, Machado Pereira AC, Nassis GP. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. Sports Medicine - Open. 2019;5:28.

Halson SL. Monitoring training load to understand fatigue in athletes. Sports Medicine. 2014;44(Suppl 2):S139–S147.

Huang E. Artificial intelligence and wearable technology in athlete load management and injury prediction: a systematic literature review. BMC Sports Science, Medicine and Rehabilitation. 2026. doi:10.1186/s13102-026-02016-w.

Leckey C, van Dyk N, Doherty C, Lawlor A, Delahunt E. Machine learning approaches to injury risk prediction in sport: a scoping review with evidence synthesis. British Journal of Sports Medicine. 2025. doi:10.1136/bjsports-2024-108576.

Lou S, Zhu H, He Z, Ye B, Sun G, Shen C, et al. The application of machine learning in the prediction of sports injuries: systematic review and meta-analysis. Scientific Reports. 2026;16:27254. doi:10.1038/s41598-026-59571-y.

Moradi M, et al. Predicting ACL injury risk in athletes: a systematic review of machine learning-based models. The Knee. 2026. doi:10.1016/j.knee.2026.104581.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71.

Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160. doi:10.1136/bmj.n160.

Rossi A, Pappalardo L, Cintia P, et al. Effective injury forecasting in soccer with GPS training data and machine learning. PLoS ONE. 2018;13:e0201264.

Van Eetvelde H, Mendonça LD, Ley C, Seil R, Tischer T. Machine learning methods in sport injury prediction and prevention: a systematic review. Journal of Experimental Orthopaedics. 2021;8:27. doi:10.1186/s40634-021-00346-x.

Vanrenterghem J, Nedergaard NJ, Robinson MA, Drust B. Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways. Sports Medicine. 2017;47:2135–2142.

Windt J, Gabbett TJ. How do training and competition workloads relate to injury? The workload-injury relationship and its implications for injury prevention. British Journal of Sports Medicine. 2019;53:1395–1397.