Original Article

doi: 10.16926/par.2026.14.03

Artificial Intelligence in Injury Prediction: A Review of Ethical Practices and Technical Standards in Elite Football

Zbigniew Waśkiewicz [1ABCD], **Tomasz Grzywacz** [2BD], **Kajetan Słomka** [3ACD], **Jacek Jurkojć** [4ADE], **Grzegorz Juras** [3ABDE]

[1] Department of Sport Management, Jerzy Kukuczka Academy of Physical Education, Katowice, Poland,
[2] Faculty of Physical Culture, Gdansk University of Physical Education and Sport, Gdańsk, Poland,
[3] Department of Motor Behavior, Jerzy Kukuczka Academy of Physical Education, Katowice, Poland,
[4] Department of Biomechatronics, Faculty of Biomedical Engineering, Silesian University of Technology, Gliwice, Poland

Authors' Contribution: A – Study Design, B – Data Collection, C – Statistical Analysis, D – Manuscript Preparation, E – Funds Collection

Abstract: The use of artificial intelligence (AI) for injury prediction in professional soccer has expanded rapidly, offering promising tools for performance optimization and medical decision-making. However, integrating AI into athlete health monitoring raises critical ethical concerns, particularly regarding algorithmic bias, model transparency, and player autonomy. This systematic review assesses the existing literature on AI-based injury prediction systems in professional soccer, with a focus on ethical dimensions, including bias mitigation, explainability, stakeholder rights, and governance frameworks. A systematic search was conducted in January 2025 across the Semantic Scholar database, yielding 497 records. Following PRISMA 2020 guidelines, 14 studies were included in the final synthesis after duplicate removal and full-text eligibility assessment. Extracted data focused on AI model types, bias identification, mitigation strategies, ethical safeguards, and stakeholder considerations. The included studies employed various AI techniques, including decision trees, XGBoost, support vector machines, and neural networks. While performance metrics were frequently reported, only six studies addressed ethical concerns. Common challenges included class imbalance, limited generalizability, and lack of transparency. Few studies discussed data ownership, consent, or the downstream implications of algorithmic decisions on player welfare. The current literature demonstrates a growing technical focus on AI-driven injury prediction but lacks robust integration of ethical and bias-related considerations. Future development of such systems should prioritize fairness audits, athlete-centered governance, and explainable AI to ensure responsible and equitable application in elite sports contexts.

Keywords: Artificial intelligence, injury prediction, professional soccer, algorithmic bias, explainable AI, athlete autonomy, ethics in sports technology

Corresponding author: Kajetan Słomka, k.slomka@awf.katowice.pl

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Recevied: 05.07.2025; Accepted: 15.09.2025; Published online: 7.01.2026

Citation: Citation: Waśkiewicz Z, Grzywacz T, Słomka K, Jurkojć J, Juras G. Artificial Intelligence in Injury Prediction: A Review of Ethical Practices and Technical Standards in Elite Football. Phys Act Rev 2026; 14(1): 26-41. doi: 10.16926/par.2026.14.03

INTRODUCTION

The increasing adoption of artificial intelligence (AI) in professional sports has transformed athlete monitoring, training optimization, and injury prevention strategies [1,2]. AI-based systems for injury prediction are gaining traction in professional soccer, where player availability, health outcomes, and performance continuity are critical to team success and individual careers [3,4]. Machine learning algorithms trained on biometric, physiological, and workload data are now used to forecast injury risk with growing accuracy and computational sophistication [5,6].

While the technical promise of these systems is well-recognized, their ethical and social implications remain underexplored [7,8]. Injury prediction models influence real-time training loads, player selection, medical intervention, and contract negotiation decisions [9,10,11]. Yet these systems are often developed and deployed without considering player consent, data ownership, transparency, or accountability for predictive errors [12,13]. Furthermore, the risk of algorithmic bias—such as models performing less accurately across subgroups defined by age, position, or injury history—raises significant fairness concerns [14–16].

The sports science and AI ethics works of literature increasingly call for explainable and athlete-centered approaches to designing and governing predictive technologies [17,18]. However, little is known about how current injury prediction systems in elite football address algorithmic fairness, model interpretability, and player rights issues [3,19]. Without adequate safeguards, AI applications may reinforce structural inequities, reduce trust in medical technologies, or disempower athletes from participating in decisions about their bodies [7,8].

This review critically examines the state of research on AI-driven injury prediction in professional soccer with a specific emphasis on ethical dimensions. We aim to identify how existing models address—or fail to address—bias detection, explainability, informed consent, and governance. By synthesizing findings from 14 empirical and conceptual studies, we aim to highlight gaps in ethical oversight and propose directions for more equitable, transparent, and accountable AI applications in sports medicine.

MATERIAL AND METHODS

Search Strategy and Data Sources

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines [20]. A comprehensive search was conducted in January 2025 using the Semantic Scholar platform, which functions as a large-scale aggregator of peer-reviewed literature (e.g., PubMed, IEEE, SpringerLink). To avoid terminological ambiguity, we classify the present work as a review rather than a systematic review, while maintaining full adherence to PRISMA selection and reporting standards. The search query used was:

("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "neural networks")

AND ("injury prediction" OR "injury forecasting" OR "injury risk" OR "injury model" OR "injury detection")

AND ("soccer" OR "football" OR "professional football" OR "elite soccer")

AND ("ethics" OR "ethical" OR "bias" OR "algorithmic bias" OR "fairness" OR "transparency" OR "explainability" OR "autonomy" OR "consent")

The initial search retrieved 497 records, and the metadata was exported into a screening tool for eligibility analysis. No limits were imposed on publication date or geographic region. Only English-language papers were included.

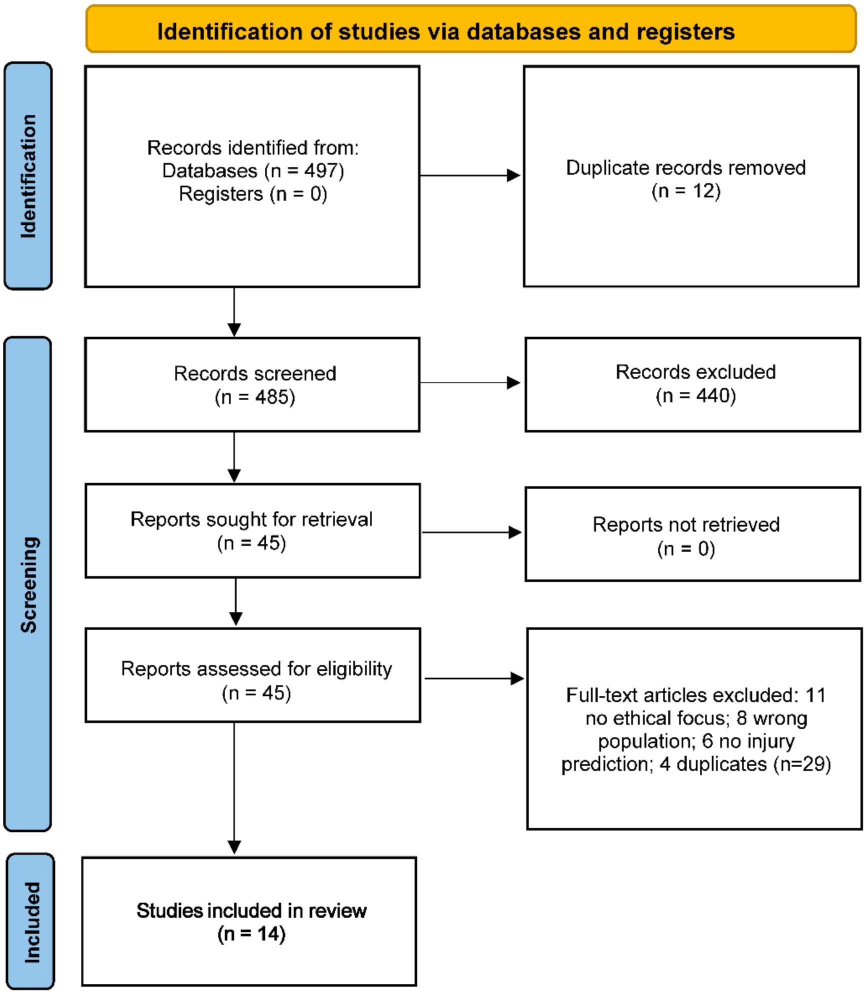

Figure 1. PRISMA 2020 flow diagram of the study selection process.

Inclusion and Exclusion Criteria

Papers were included if they met all the following criteria:

1. Focused on AI-driven injury prediction systems in professional-level soccer.
2. Included some discussion of ethical implications, algorithmic bias, and/or model interpretability or ethical/bias considerations.
3. Empirical study, case study, conceptual paper, or systematic review with methodological grounding.
4. Provided sufficient detail on the AI model(s) used, data collection processes, or implementation context.

Studies were excluded if:

1. They addressed injury prediction without AI-based methods.
2. They focused solely on youth/amateur sports.
3. No ethical or bias-related discussion was presented.
4. They were editorials, commentaries, or lacked methodological clarity.

Study Selection and Screening Process

After removing 12 duplicate records, two independent reviewers screened 485 titles and abstracts. A total of 440 studies were excluded for not meeting the inclusion criteria. The remaining 45 articles underwent full-text review, resulting in the final inclusion of 14 studies relevant to the ethical and algorithmic dimensions of AI-driven injury prediction in professional soccer.

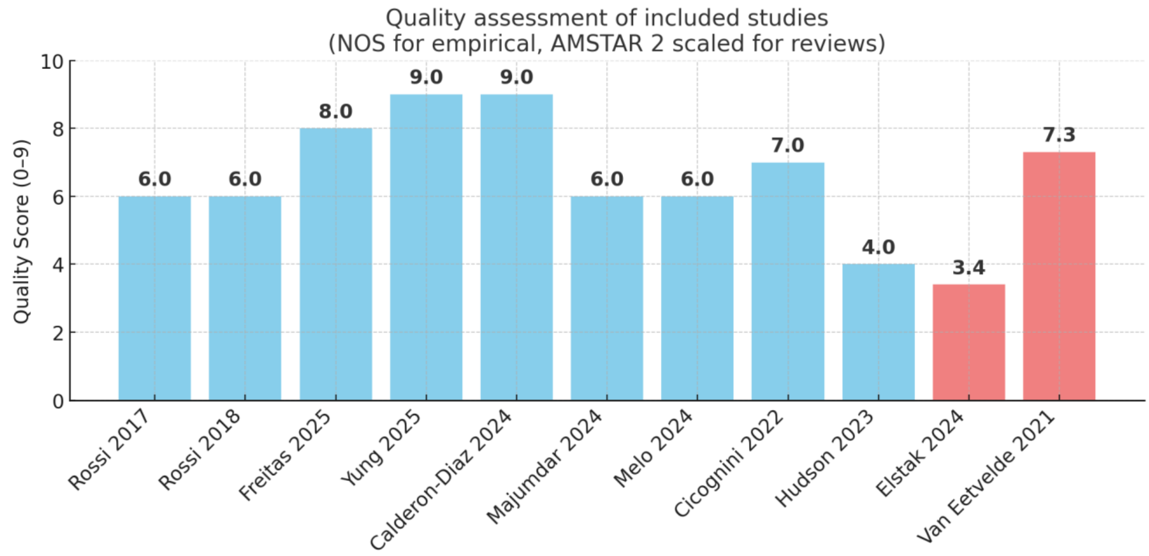

Figure 2. Quality assessment of included studies. Empirical studies were evaluated using the Newcastle–Ottawa Scale (NOS, blue bars; range 0–9), while systematic reviews were assessed with AMSTAR 2 (red bars; scores rescaled to 0–9 for comparability)

Data Extraction

From each included study, we extracted structured data on the following dimensions:

a. Study focuses and population (e.g., type of injury, player sample size, team level)
b. AI models and techniques (e.g., decision trees, XGBoost, CNNs, SHAP explainability)
c. Ethical framework presence (e.g., informed consent, autonomy, data ownership)
d. Identified biases (e.g., class imbalance, overfitting, generalizability issues)
e. Bias mitigation strategies (e.g., PCA, undersampling, interpretable models)
f. Discussion of player rights and autonomy
g. Model transparency and explainability
h. Implementation challenges and stakeholder impact

Quality Assessment

In this review, a quality assessment of the included studies was conducted using standardized tools appropriate to the study type. Empirical studies were evaluated with the Newcastle–Ottawa Scale (NOS), while systematic reviews were assessed with the AMSTAR 2 tool (Figure 2). This procedure enabled a structured and comparable appraisal of methodological rigor, covering criteria such as participant selection, control of confounding factors, outcome measurement, and transparency of reporting. Incorporating quality assessment was an essential step in the methodological process, ensuring consistency of analysis and supporting a more critical interpretation of the available evidence.

Inclusion of Real-World AI Systems

To complement the systematic review of peer-reviewed academic studies, we conducted a parallel exploratory analysis of real-world AI systems currently deployed in elite professional football. The aim was to assess how injury prediction tools are implemented in practice, particularly in terms of transparency, validation, explainability, and ethical safeguards. We identified prominent systems such as Zone7, Kitman Labs' Risk Advisor, and FC Bayern Munich's in-house AI platform through the purposive sampling of grey literature, including:

a. Official vendor websites and product documentation
b. Interviews and statements from club personnel

c. Industry publications and sports technology media
d. Public case studies and press releases

These sources were included based on the following criteria:

a. The system must be actively deployed in professional football settings
b. The system must involve machine learning or AI-based injury risk forecasting
c. The system must have sufficient publicly available information to allow structured analysis

Although these systems do not meet the formal inclusion criteria for systematic review (e.g., peer-reviewed evidence), they represent the de facto state of applied AI in sports and offer a critical context for interpreting the academic literature. We extracted information on system design, data inputs, decision-making protocols, explainability, validation claims, and ethical framing. Data were synthesized thematically and organized in comparative tables (see Tables 3 and 4) to highlight contrasts with the academic domain. As these are commercial or internal systems, the analysis is limited to publicly disclosed information and should be interpreted accordingly.

RESULTS

This section presents a detailed synthesis of findings from fourteen studies that met the eligibility criteria for this systematic review, supplemented by a comparative analysis of leading real-world AI systems currently deployed in elite professional football. The aim is twofold: first, to examine how the academic literature conceptualizes, implements, and evaluates AI-based injury prediction systems, particularly concerning issues of transparency, bias, and ethical governance; and second, to contextualize these findings by analyzing how AI is applied in practice by clubs, commercial vendors, and in-house analytics teams.

The results are presented in two stages. The first synthesizes evidence from peer-reviewed studies, detailing their methodological characteristics, AI model architectures, data types, bias mitigation approaches, and attention to ethical dimensions. The second provides an overview of commercial and proprietary systems—such as Zone7, Kitman Labs, and FC Bayern Munich's in-house AI—that are actively used in top-tier clubs and often influence industry standards.

We align these findings where possible to facilitate direct comparison of research models and applied systems across key domains, including data scope, explainability, validation, and decision-making protocols. This dual-level approach enables a more nuanced assessment of the current state of injury prediction in professional football, highlighting the gap between technological capability and ethical maturity.

Characteristics of Included Academic Studies

The fifteen studies included in this review were published between 2017 and 2025, spanning a range of empirical, technical, and conceptual contributions (Table 1). Two publications [23, 26] reported analyses of the same cohort of 284 players from 16 clubs. To avoid duplication, only the more comprehensive study [26] was included, resulting in a final total of 14 studies.

All studies focused explicitly on injury prediction or injury risk forecasting in the context of professional male soccer players, with no studies addressing female cohorts or amateur athletes. While most studies were confined to single-team datasets or limited league cohorts, several attempted to generalize findings across multiple seasons or populations. Across the corpus, studies employed various supervised machine learning techniques, often using retrospective datasets of player health, workload, and physical performance metrics. Typical data sources included GPS-based load tracking, training session logs, biomechanical metrics, and injury event records drawn from club databases. However, few studies have incorporated psychological, sociocultural, or contextual data, such as athlete-reported wellness, mood, or external life stressors, which are often relevant in practitioner experience and anecdotal reports.

Sample sizes varied significantly. The smallest study analyzed data from 26 players, while the largest examined 3,374 player-seasons using epidemiological league-wide records. Most studies were based on a single club's dataset, with limited geographic diversity, although one study drew from teams in three countries. Notably, only one study [4] explicitly reported compliance with ethical research protocols, such as obtaining informed consent and ensuring anonymization.

The comparative analysis of the included studies reveals several noteworthy patterns. First, most models were developed using data from a single team or club, which raises concerns about the lack of external validation and limited generalizability of the findings across different competitive contexts or athlete populations. Second, there was significant diversity in the machine learning approaches adopted, with ensemble methods such as XGBoost and AdaBoost, tree-based classifiers, and various forms of neural networks all prominently featured. Additionally, Yung et al. [21] employed a Bayesian network trained on league-wide epidemiological records to classify time to return to sport, illustrating the use of probabilistic graphical models in this domain.

This methodological variety reflects the experimental nature of the field but also complicates comparisons of model performance and reliability. Relatedly, Pappalardo et al. [22] proposed the PlayeRank machine-learning framework for player evaluation, which, although not a direct predictor of injury, contextualizes risk through performance tracking and workload signals.

Third, the data sources used in these studies were overwhelmingly based on GPS-derived training load metrics, with relatively few studies incorporating biomechanical assessments, video-derived data, or epidemiological injury records. For example, Navarro et al. [23] combined biomechanical testing and sprint profiling with decision-tree and logistic-regression models to estimate hamstring injury risk across multi-club cohorts.

This narrow data scope may limit the holistic evaluation of injury risk and overlook potentially meaningful physiological or contextual factors. Finally, attention to ethical procedures was generally poor across the literature. Among the included studies, six explicitly or implicitly engaged with ethical concerns. Freitas et al. [4] reported obtaining informed consent and ensuring anonymization of GPS-derived player data, directly addressing data protection requirements. Procopiou and Piki [17] provided a conceptual analysis of governance and fairness in AI-driven systems, arguing that opaque injury prediction tools risk undermining athlete autonomy. Several empirical works also incorporated safeguards that align with ethical standards. Hudson et al. [24] emphasized transparency by applying explainable multi-modal models, allowing practitioners to interrogate which features drove injury risk classifications. Similarly, Melo et al. [6] used SHAP-based approaches to evaluate model fairness and mitigate bias, thereby linking interpretability to ethical accountability. From a broader perspective, Elstak et al. [3] reviewed AI applications in football and highlighted issues of bias and governance as central limitations for practical deployment. Finally, Majumdar et al. [25] emphasized the importance of open science practices and reproducibility in football injury prediction research, explicitly situating transparency as a core ethical principle. Collectively, these six contributions show that while ethical safeguards are not yet consistently implemented, emerging work demonstrates practical pathways for integrating transparency, fairness, and data governance into AI-based injury prediction.

Bias Mitigation, Explainability, and Ethical Integration in Academic Models

In addition to methodological characteristics, we examined how the included studies addressed core ethical and technical issues relevant to the responsible deployment of AI, namely, algorithmic bias, model transparency, and the presence of athlete-centered governance structures (Table 2). While injury prediction inherently involves imbalanced datasets, given that injuries are relatively rare, only a subset of studies implemented explicit bias mitigation strategies. Freitas et al. [4] employed stratified cross-validation to improve generalization across classes. In contrast, Melo et al. [6] utilized a combination of undersampling and principal component analysis (PCA) to address concerns related to imbalance and dimensionality. Other studies did not address mitigation efforts or fairness metrics, despite working with highly skewed data.

Model explainability was inconsistently addressed. Some studies, such as those by Hudson et al. [9] and Majumdar et al. [25], implemented SHAP (Shapley Additive Explanations) or interpretable local modeling techniques, enabling the identification of features most influential in driving model outputs. In a biomechanical context, Calderón-Díaz et al. [26] applied explainable ML and used SHAP analyses to identify muscle-injury risk factors in professional players.

In contrast, several studies relying on complex architectures, such as convolutional or artificial neural networks, did not provide a mechanism for model interpretability, instead focusing solely on predictive accuracy. Notably, explainability was framed almost exclusively as a technical feature—benefiting researchers or sports science staff—rather than as a communicative or ethical obligation to the athletes whose data was being analyzed. No study included athlete-facing interfaces or mechanisms for feedback and contestability.

The reviewed studies used no standardized framework to assess fairness, such as evaluating whether models perform equally well across subgroups (e.g., by position, age, or injury history). Moreover, none of the empirical studies discussed accountability structures in cases where AI-generated injury forecasts might lead to incorrect human decisions, such as benching a healthy player or ignoring a legitimate injury risk. The lack of integration between technical design and ethical foresight reflects a broader pattern in sports analytics: while AI systems advance in predictive power, they are rarely developed with concurrent attention to transparency, equity, or athlete rights. This becomes particularly evident when these academic models are compared with the approaches used in real-world deployments, which we turn to in the following section.

Real-World AI Systems in Professional Soccer

While academic research on AI-driven injury prediction is confined mainly to experimental models and retrospective datasets, professional football clubs already deploy AI systems in high-stakes, real-time environments (Table 3). These proprietary systems are shaping the norms of injury risk management, often without the benefit—or burden—of peer review. To complement the findings from the academic literature, we analyzed three prominent systems that have seen widespread adoption or institutional investment: Zone7 [27], Kitman Labs' Risk Advisor [28,29], and FC Bayern Munich's in-house AI platform [30]. These systems vary in architecture, user interface, transparency, and integration into club workflows. Yet, they all share the core objective of identifying players at elevated risk of injury and supporting preventive interventions.

What distinguishes these real-world systems from academic models is not necessarily their predictive architecture, which often remains undisclosed, but rather the contextual complexity in which they operate. These platforms are embedded into the daily operations of professional clubs, requiring user trust, data integration across departments, and responsiveness to medical and tactical priorities. Their performance is evaluated based on statistical accuracy, practical utility, interpretability, and ethical acceptability within high-performance sporting environments.

Table 1. Summary of Included Academic Studies on AI-Based Injury Prediction in Professional Soccer

Author	Country / League	Population / N / Seasons / Injuries	Data source(s)	Methods (AI/ML)	Outcome	Validation	XAI / Ethics
Rossi et al. (2017)	Italy (Serie A, 1 club)	Pro male; N=26; 1 season; 23 injuries	GPS workload	Decision tree, Random Forest, Logistic Regression	Injury prediction (training/match)	Internal CV	—
Rossi et al. (2018)	Italy (Serie A, 1 club)	Pro male; N=26; 1 season; 23 injuries	GPS workload	Decision tree + ADASYN	Injury forecast (next session)	Internal resampling (2-fold repeated)	Case-based rules (interpretable)
Freitas et al. (2025)	Portugal (Primeira Liga, 1 club)	Pro male; N=34; 1 season; 38 non-contact injuries	GPS Catapult + player descriptors	SVM, FNN, AdaBoost; mRMR	Daily injury risk	Internal CV	Ethics council data request noted
Yung et al. (2025)	Germany (Bundesliga)	Pro male; 7 seasons; 3374 player-seasons; 6143 injuries	Public epidemiological/media data	Bayesian Network	RTS classification; injury severity	Temporal split (train/test)	BN is transparent; consent not applicable
Calderón-Díaz et al. (2024)	Japan, France, Finland (16 clubs)	Pro male; N=284; 1 season; 47 hamstring injuries (u 38 zawodników)	Biometrics + sprint profile + history of injuries	Navarro: Logistic Regression, Decision Tree; Calderón-Díaz: XGBoost, SHAP	Hamstring injury prediction	Nested CV (2000 runs)	Yes (SHAP, interpretable models)
Hudson et al. (2023)	Mixed sports incl. football	Dataset not specified (proof-of-concept)	Multi-modal sports data	Local & global explainable models	Injury risk modelling	Internal only	Yes (XAI local explanations)
Melo et al. (2024)	Brazil (pro club)	Pro male; N=52; 2 seasons; 63 injuries	GPS microcycle features	Data-centric AI + SHAP	Injury prediction (microcycle)	Internal	Yes (SHAP explainability)
Majumdar et al. (2024)	UK clubs	Pro male; N=120; 3 seasons; 146 injuries	GPS workload	ML ensemble models	Injury occurrence	Internal	—
Chen & Sirisena (2024)	Asia (multi-sport)	Not applicable (framework paper)	Wearables + tracker data	AI-SPOT system	Injury risk detection & training optimization	—	Mentions governance
Cicognini et al. (2022)	Argentina (Belgrano de Córdoba)	Pro male; 10 years (2010–2019); ~80,000 sessions; 300 injuries train + 39 test	GPS + RPE metrics	Logistic Regression, Decision Tree, RF, GBM	Non-contact injury prediction	Temporal split (2013–18 train, 2019 test)	SHAP explainability
Elstak et al. (2024)	Multi-country	190 studies reviewed	Literature	—	AI in football codes	—	Bias & governance discussed
Van Eetvelde et al. (2021)	Multi-sport	28 studies reviewed	Literature	—	ML in injury prediction	—	Notes bias & lack of validation
Pappalardo et al. (2019)	Italy (Serie A 2004–16)	21M match events	Match event + tracking data	PlayeRank ML framework	Player ranking (context for workload)	—	Governance context only
Procopiou & Piki (2023)	Europe	Conceptual (no dataset)	Conceptual /XAI	Framework for sport AI	—	—	Governance, fairness
West, Shrier, Impellizzeri et al. (2025)	International	Conceptual (no dataset)	Training load critique	—	—	—	Ethics of misuse, governance

Table 2. Ethical and Technical Integration Across Academic Studies

Study	Bias Mitigation Reported	Explainability Tools	Ethical Safeguards Mentioned	Athlete Autonomy Considered
Freitas et al. (2025)	Stratified cross-validation	None (performance-focused)	Yes (consent, anonymization)	No
Melo et al. (2024)	Undersampling, PCA	SHAP, feature importance	No	No
Majumdar et al. (2024)	None	SHAP	No	No
Rossi et al. (2017/2018)	None	Interpretable decision tree rules	No	No
Hudson et al. (2023)	None	Local modeling, XAI	No	No
Calderón-Díaz et al. (2024)	None	SHAP, feature importance (XGBoost)	No	No
Yung et al. (2025)	None	None	No	No
Cicognini et al. (2022)	Temporal validation split	SHAP (GBM interpretation)	No	No
Pappalardo et al. (2019)	None	None (framework)	No	No
Procopiou & Piki (2023)	Conceptual discussion of fairness	Conceptual transparency, XAI principles	Yes (full ethical design framework)	Yes
Elstak et al. (2024)	Not applicable (systematic review)	Not applicable	Notes bias/governance but no safeguards	No
Van Eetvelde et al. (2021)	Not applicable (systematic review)	Not applicable	No	No
Chen & Sirisena (2024)	Not applicable (framework paper)	None	Mentions governance	No
West et al. (2025)	Conceptual critique of misuse	None	Discusses ethical misuse of metrics	Yes (via athlete rights framing)

Zone7 exemplifies the integration of multi-club data and machine learning to provide each player with daily or weekly risk stratification. Clubs such as Getafe CF and Rangers FC have publicly credited Zone7 with reducing injury rates, although these outcomes are reported through internal audits rather than scientific publications. The system generates alerts but leaves intervention decisions to coaching or medical staff, positioning itself as a decision-support tool rather than an autonomous actor. However, the proprietary nature of the model—combined with the absence of peer-reviewed validation—limits independent assessment of fairness, subgroup sensitivity, or false favorable rates.

Kitman Labs' Risk Advisor platform adopts a more explainability-focused approach, allowing clubs to trace which metrics contribute to each player's risk flag. Unlike Zone7, Kitman encourages users to configure their alert thresholds and data inputs, offering a more transparent and collaborative AI interface. Yet, despite a stronger emphasis on interpretability and trust, Kitman lacks independent evaluation of model efficacy or fairness, and its algorithmic core remains opaque.

FC Bayern Munich's in-house AI system represents a different archetype: a fully bespoke, non-commercial solution built on extensive internal data infrastructure. This system reportedly utilizes computer vision and deep learning on training and match footage to identify subtle deviations in movement patterns that may signal an increased risk of developing injury. Coaches receive real-time feedback and adjust workloads accordingly. While Bayern's system is widely cited as a benchmark for technological integration in elite sports, it is not subject to external audit or reproducibility, and no technical details have been published in the academic domain.

These real-world systems present a compelling counterpoint to the academic literature (Table 4). While academic models often emphasize algorithmic novelty and internal accuracy metrics, commercial and club-deployed systems foreground practical utility, user trust, and operational integration. Yet, both domains suffer from similar shortcomings: limited transparency, lack of subgroup fairness testing, and absence of regulatory oversight. This convergence of strengths and limitations will be analyzed further in the Discussion section.

This synthesis reveals an apparent disconnect between the academic and applied domains. Academic studies prioritize methodological development but lack generalizability, user feedback loops, and ethical frameworks. Real-world systems, by contrast, are embedded in professional workflows and cater to the needs of practitioners; however, they often sacrifice scientific transparency and standardized evaluation in the process. Notably, neither sphere currently integrates formal fairness audits, athlete-centered governance, or third-party regulation, despite the increasing influence of AI tools on player health and career trajectories. This dual deficiency—technical opacity in practice and ethical immaturity in research—represents a central challenge in the responsible development of AI in sports, which we will explore in the following Discussion section.

Table 3. Comparative Overview of Real-World AI Injury Prediction Systems in Professional Football

System	Clubs / Leagues Using It	Primary Data Sources	Explainability	Decision-Making Role	Transparency	Validation / Outcomes
Zone7	EPL, La Liga, MLS, Getafe CF, Rangers FC	GPS metrics, biometric data, tracking data	Moderate – includes contributing factors, no full model access	Human-in-the-loop	Proprietary "black box"	Claimed 72–75% sensitivity; Getafe CF and Rangers FC reported injury reductions (up to 52%)
Kitman Labs – Risk Advisor	Multiple elite clubs (undisclosed), global adoption	Training loads, wellness surveys, physio notes	High – configurable dashboards, risk driver breakdowns	Human-in-the-loop	Proprietary but user-configurable	Anecdotal improvements in player availability; no peer-reviewed outcomes published
Bayern Munich In-House AI	Internal to FC Bayern Munich; model for PSG, Liverpool	Computer vision, video analysis, biomechanical monitoring	Real-time flagging of biomechanical anomalies, internal-only	Human-in-the-loop	Not externally disclosed	Club reports suggest reduced injury incidence; no external validation or transparency available

Table 4. Synthesis Comparison of Academic Studies and Real-World AI Systems for Injury Prediction in Professional Soccer

Dimension	Academic Models (n = 14)	Real-World Systems
Validation	Typically, internal only (train/test split on same dataset); no external or cross-club validation in any empirical study	Multi-environment deployment (e.g., Zone7 across leagues); Bayern and Kitman validated internally but not peer-reviewed
Data Scope	Primarily GPS data, training loads; limited use of video, biomechanics, or contextual variables	Multimodal inputs including GPS, biomechanics, computer vision, wellness data, and video analytics (e.g., Bayern's AI)
Sample Size	Small, often single-team (26–284 players); one large epidemiological study (3,374 player seasons)	Aggregated across clubs/seasons in Zone7; full squad coverage in Kitman and Bayern; unknown but presumed larger than academic data
Explainability	Inconsistent; few used SHAP or decision trees; rarely athlete-facing	Stronger emphasis (e.g., Kitman: explainable dashboards; Zone7: contributing factors shown); Bayern provides real-time visual cues
Ethical Safeguards	Only one study (Freitas et al. 2025) reported informed consent; no discussion of opt-out, rights to explanation, or contestability	Human-in-the-loop approach emphasized; informal norms around coach accountability and data protection; no formal ethical standards
Fairness Analysis	Absent in all studies; no subgroup performance tested	Absent in vendor reports; unknown whether subgroup biases (e.g., by position, injury history) are measured
Transparency	Moderate: Methods described but code rarely shared; models often not reproducible	Low: Proprietary or internal systems; algorithms not publicly available; performance metrics reported only in press releases
Decision Autonomy	Not discussed explicitly; all assume human decision-making	Emphasized: all systems used as decision-support, not autonomous agents
User Integration	Designed for academic analysis or technical staff; not embedded in daily workflows	Embedded into daily club operations; coach and medical interfaces are core to Zone7 and Kitman adoption
Regulatory Oversight	None; no mention of GDPR, medical device standards, or sport-specific governance	Nonformal; compliance with general data protection (e.g., GDPR) assumed but no domain-specific regulation

DISCUSSION

This systematic review examined the state of AI-driven injury prediction in professional football through two lenses: the academic literature, comprising fifteen empirical and conceptual studies, and real-world systems currently deployed by elite clubs. These perspectives provide a nuanced understanding of the technological and ethical development of injury prediction in sports. They reveal both the promise of predictive analytics and the scientific, procedural, and ethical gaps that must be addressed to ensure these tools are deployed responsibly.

Academic research: promise and persistent limitations

The academic literature demonstrates significant technical ambition, with diverse applications of decision trees, ensemble models, neural networks, and Bayesian approaches to GPS-derived data, biomechanical measures, and injury histories. Yet these models remain constrained by methodological insularity: most rely on single-club datasets, rarely attempt external validation, and seldom assess fairness across subgroups defined by age, position, or previous injury history. While some studies experimented with explainability methods such as SHAP, these were primarily technical demonstrations aimed at researchers rather than communication tools for athletes. Ethical safeguards were generally minimal, with informed consent or anonymization explicitly reported in only one case. As a result, the academic evidence base remains fragmented, technically promising but ethically immature, and insufficiently prepared for translation into practice.

Commercial systems: pragmatic integration without transparency

Commercial platforms such as Zone7, Kitman Labs, and FC Bayern Munich's in-house AI illustrate that predictive systems have already penetrated elite football environments. Their integration into daily training and medical workflows demonstrates a high degree of practical feasibility. Features such as configurable dashboards, risk alerts, and real-time video feedback make these tools attractive to practitioners. However, this pragmatic innovation comes at the expense of transparency. Proprietary algorithms remain inaccessible, performance claims are often based on internal audits rather than peer-reviewed validation, and no independent evidence is available on subgroup fairness. In contrast to the academic literature, these systems are deeply embedded in professional practice, but they operate as "black boxes" in which player rights, contestability, and accountability mechanisms are absent.

Shared gaps and systemic implications

Despite their differences, academic and commercial approaches converge on one point: the absence of structured, athlete-centered ethical frameworks. Neither sphere incorporates systematic fairness audits, robust governance protocols, or clear accountability when predictions are wrong. This gap is not only technical but systemic, reflecting a lack of regulatory oversight comparable to that already present in other domains such as digital health or medical AI. In medicine, the introduction of regulatory pathways (e.g., FDA and EMA approvals, medical device directives) has forced developers to adopt standards of transparency, validation, and patient protection. Elite football currently lacks such frameworks, leaving clubs to navigate ethical responsibilities on their own, while it is known that decision-making support systems in sports can increase the efficiency and effectiveness of the development of players individually as well as in the future of entire teams [31]. Moreover, economic and legal realities limit what can be realistically implemented. Clubs operate under intense financial pressure, and additional oversight structures may be deprioritized in favor of short-term performance objectives. Without viable funding mechanisms and regulatory incentives, proposals such as ethical audits or inter-club data validation risk remaining aspirational. The challenge, therefore, is to create an ecosystem where rigorous methodological standards and enforceable ethical safeguards evolve in tandem, ensuring that AI-driven injury prediction serves not only clubs but also the athletes whose health and careers are directly impacted.

Importantly, the feasibility of these recommendations must be assessed in the light of economic and legal realities. Clubs often operate under budgetary pressure, and compliance with additional oversight frameworks may be deprioritized. Without viable funding mechanisms and regulatory incentives, ethical guidelines risk remaining aspirational.

Recommendations for Future Research and Practice

To close these gaps, we recommend a series of practical and conceptual steps:

1. *Mandate External Validation:* Future academic studies should test model generalizability using external datasets from different teams, leagues, or cohorts. Journals and conferences should require this as part of standard reporting.
2. *Integrate Fairness Audits*: Researchers and vendors must evaluate model performance across key subgroups (e.g., age, position, injury history) and report disparities transparently.
3. *Build Athlete-Centered Interfaces:* Explainability must extend beyond staff dashboards. Players should be able to view, understand, and respond to injury risk assessments in a psychologically supportive and context-aware manner.
4. *Formalize Ethical Safeguards:* Clubs, vendors, and research institutions should establish standardized protocols for obtaining consent, accessing data, exercising opt-out rights, and handling errors. These frameworks should be integrated into collective bargaining agreements and athlete welfare policies.
5. *Pursue Regulatory and Scientific Oversight:* Sports governing bodies, medical ethics boards, and AI governance experts must collaborate to establish minimal ethical and technical standards for predictive systems in elite sports. This may include certification schemes, independent audits, and public reporting requirements.

CONCLUSION

AI-based injury prediction systems in professional football are transitioning from experimental models to operational tools that influence training, selection, and player health decisions. While these technologies promise efficiency and preventive care gains, they are being developed and deployed within ethically fragile and scientifically fragmented ecosystems. The academic literature remains focused on model-building, with limited attention to generalizability, fairness, or player empowerment. Meanwhile, real-world systems are increasingly influential but operate with little transparency or independent validation.

Bridging this gap requires more than technical refinement—it demands a reorientation of priorities. Injury prediction tools must be evaluated by their predictive performance and capacity to support trustworthy, fair, and inclusive sports environments. This includes respecting athlete autonomy, ensuring transparency, and holding systems accountable for their recommendations and the consequences that follow. The future of AI in sports will depend not only on what these systems can predict, but also on how they are designed, governed, and utilized.

Limitations

While this review offers a comprehensive synthesis of academic and real-world developments in AI-driven injury prediction for professional football, several limitations should be acknowledged.

First, the scope of the academic literature reviewed was constrained to studies available in English and indexed in the Semantic Scholar. Although this database aggregates a wide range of publications, it may not capture all relevant works published in other languages or less-visible sports science journals. Some technical reports, theses, or non-indexed conference papers with relevant contributions may have been omitted. The majority of included studies focused on male professional cohorts, with limited data on female athletes or at the youth level.

Second, many of the included academic studies suffered from incomplete or inconsistent reporting, particularly regarding sample characteristics, data sources, model parameters, and ethical procedures. This limited our ability to perform a uniform comparison or conduct a meta-analytic assessment of model performance. In several cases, key methodological details had to be inferred or were absent, reflecting a broader lack of transparency in the domain.

Third, while the review incorporates real-world AI systems utilized in elite sports, the evidence base for these platforms is primarily derived from publicly available reports, press releases, and vendor communications. These sources are valuable, but they are inherently partial and unverified. Without peer-reviewed validation or open access to performance data, our analysis of real-world systems remains descriptive and interpretive rather than empirically confirmatory.

Fourth, the review did not perform a formal risk of bias assessment for each included study. This was due to the heterogeneity of study types (empirical, conceptual, technical prototype) and the absence of a standardized risk-of-bias tool tailored to AI-in-sport research. However, the observed variability in data quality, model validation, and ethical reporting across studies indicates the need for such an instrument in future systematic reviews.

Another limitation is the absence of protocol registration in repositories such as PROSPERO. This decision was taken because PROSPERO does not currently accommodate interdisciplinary reviews that combine AI ethics with sports science. Nevertheless, we recognize that protocol registration would have enhanced transparency and reproducibility.

Finally, this review focuses exclusively on men's professional football, reflecting the population scope of all included studies and commercial deployments. Furthermore, all academic studies and real-world systems reviewed focus solely on male professional footballers. The absence of female players and youth cohorts is a fundamental limitation that not only undermines the generalizability of current findings but also risks perpetuating gender and age-related inequalities in access to sports technology. As such, the conclusions presented here cannot be assumed to extend to women's football, youth athletes, or other sporting contexts. Given the rapid growth of AI-based performance and injury prediction tools across the sports sector, it is imperative that future research systematically includes female and youth populations, ensuring that technological development does not reinforce existing disparities but instead promotes inclusivity and equity.

Funding Statement: This research was funded by the Ministry of Science and Higher Education, Republic of Poland, grant number MNiSW/2024/DAP/225.

REFERENCES

1. Phatak AA, Wieland FG, Vempala K, Volkmar F, Memmert D. Artificial Intelligence Based Body Sensor Network Framework—Narrative Review: Proposing an End-to-End Framework Using Wearable Sensors, Real-Time Location Systems and Artificial Intelligence/Machine Learning Algorithms for Data Collection, Data Mining and Knowledge Discovery in Sports and Healthcare. Sports Med Open. 2021; 7: 372. doi: 10.1186/s40798-021-00372-0
2. Mateus N, Abade E, Coutinho D, Gómez MÁ, Peñas CL, Sampaio J. Empowering the Sports Scientist with Artificial Intelligence in Training, Performance, and Health Management. Sensors. 2025; 25(1): 139. doi: 10.3390/s25010139

3. Elstak I, Salmon P, McLean S. Artificial Intelligence Applications in the Football Codes: A Systematic Review. J Sports Sci. 2024; 42(13): 1184-1199. doi: 10.1080/02640414.2024.2383065
4. Freitas DN, Mostafa SS, Caldeira R, Santos F, Fermé E, Gouveia ÉR, Morgado-Dias F. Predicting Noncontact Injuries of Professional Football Players Using Machine Learning. PLoS One. 2025; 20(1): e0315481. doi: 10.1371/journal.pone.0315481
5. Rossi A, Pappalardo L, Cintia P, Iaia MF, Fernández J, Medina D. Effective Injury Prediction in Professional Soccer with GPS Data and Machine Learning. PLoS One. 2018; 13(7): e0201264. doi: 10.1371/journal.pone.0201264
6. Melo M, Maia M, Padrão G, Brandão D, Bezerra E, Spineti J, Giusti L, Soares J. Data-Centric AI for Predicting Non-Contact Injuries in Professional Soccer Players. In: Proceedings of the XXXIX Simpósio Brasileiro de Banco de Dados (SBBD 2024). Porto Alegre: Sociedade Brasileira de Computação; 2024. p. 167-180.
7. Karkazis K, Fishman JR. Tracking U.S. Professional Athletes: The Ethics of Biometric Technologies. Am J Bioeth. 2017; 17(1): 45-60. doi: 10.1080/15265161.2016.1251633
8. MacLean E. The Case of Tracking Athletes' Every Move: Biometrics in Professional Sports and the Legal and Ethical Implications of Data Collection. SSRN Electron J. 2020. doi: 10.2139/ssrn.3821995
9. Hudson D, Hartigh RJR den, Meerhoff LRA, Atzmueller M. Explainable Multi-Modal and Local Approaches to Modelling Injuries in Sports Data. In: Proceedings of the 2023 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE; 2023. p. 949-957.
10. Wang DY, Ding J, Sun AL, Liu SG, Jiang D, Li N, Yu JK. Artificial Intelligence Suppression as a Strategy to Mitigate Artificial Intelligence Automation Bias. J Am Med Inform Assoc. 2023; 30(10): 1684-1692. doi: 10.1093/jamia/ocad118
11. Wang Z, Velickovic P, Hennes D, Tomasev N, Prince L, Kaisers M, Bachrach Y, Elie R, Li W, Piccinini F, et al. TacticAI: An AI Assistant for Football Tactics. arXiv. 2023; arXiv:2310.10553. doi: 10.48550/arxiv.2310.10553
12. Rush C, Osborne B. Benefits and Concerns Abound, Regulations Lack in Collegiate Athlete Biometric Data Collection. J Leg Aspects Sport. 2022; 32(1): 62-94. doi: 10.18060/25479
13. Purcell RH, Rommelfanger KS. Biometric Tracking From Professional Athletes to Consumers. Am J Bioeth. 2017; 17(1): 72-74. doi: 10.1080/15265161.2016.1251652
14. Zhang J, Li J. Mitigating Bias and Error in Machine Learning to Protect Sports Data. Comput Intell Neurosci. 2022; 2022: 4777010. doi: 10.1155/2022/4777010
15. Zhang Y. Artificial Intelligence and Big Data-Based Injury Risk Assessment System for Sports Training. Mob Inf Syst. 2022; 2022: 7125462. doi: 10.1155/2022/7125462
16. Palermi S, Vecchiato M, Saglietto A, Niederseer D, Oxborough D, Ortega-Martorell S, Olier I, Castelletti S, Baggish A, Maffessanti F, et al. Unlocking the Potential of Artificial Intelligence in Sports Cardiology: Does It Have a Role in Evaluating Athlete's Heart? Eur J Prev Cardiol. 2024; 31(4): 470-482. doi: 10.1093/eurjpc/zwae008
17. Procopiou A, Piki A. The 12th Player: Explainable Artificial Intelligence (XAI) in Football: Conceptualisation, Applications, Challenges and Future Directions. In: Proceedings of the 19th International Conference on Web Information Systems and Technologies. SCITEPRESS; 2023. p. 213-220.
18. Angus AC, Sirisena D. AI-SPOT: A Novel Artificial Intelligence-Enabled Sport Optimization Tool to Enhance Performance and Prevent Injury in Elite Footballers. Glob Med Health Commun. 2024; 12(1): 86-92. doi: 10.29313/gmhc.v12i1.12788
19. Bullock GS, Mylott J, Hughes T, Nicholson KF, Riley RD, Collins GS. 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. Sports Med. 2022; 52(10): 2469-2482. doi: 10.1007/s40279-022-01698-9
20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. BMJ. 2021; 372: n71. doi: 10.1136/bmj.n71
21. Yung KKY, Wu PPY, aus der Fünten K, Hecksteden A, Meyer T. Using a Bayesian Network to Classify Time to Return to Sport Based on Football Injury Epidemiological Data. PLoS One. 2025; 20(1): e0314184. doi: 10.1371/journal.pone.0314184
22. Pappalardo L, Cintia P, Ferragina P, Massucco E, Pedreschi D, Giannotti F. PlayeRank: Data-Driven Performance Evaluation and Player Ranking in Soccer via a Machine Learning Approach. ACM Trans Intell Syst Technol. 2019; 10(5): 1-27. doi: 10.1145/3343172

23. Navarro L, Dandrieux PE, Hollander K, Edouard P. Digitalization in Professional Football: An Opportunity to Estimate Injury Risk. In: IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing; 2022. p. 366-375.
24. Hudson D, den Hartigh RJR, Meerhoff LRA, Atzmueller M. Explainable Multi-Modal and Local Approaches to Modelling Injuries in Sports Data. In: Proceedings of the 2023 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE; 2023. p. 949-957.
25. Majumdar A, Bakirov R, Hodges D, McCullagh S, Rees T. A Multi-Season Machine Learning Approach to Examine the Training Load and Injury Relationship in Professional Soccer. J Sports Anal. 2024; 10(1): 47-65. doi: 10.3233/jsa-240718
26. Calderón-Díaz M, Silvestre Aguirre R, Vásconez JP, Yáñez R, Roby M, Querales M, Salas R. Explainable Machine Learning Techniques to Predict Muscle Injuries in Professional Soccer Players through Biomechanical Analysis. Sensors. 2024; 24(1): 119. doi: 10.3390/s24010119
27. Zone7. Validation Study: Injury Risk Forecasting with Zone7 AI. Technical Report. 2022.
28. Kitman Labs. Risk Advisor: Data-Driven Injury Risk Assessment. Technical Report. 2022.
29. Kitman Labs. The Intelligence Platform: Modernizing Performance and Medical Operations in Sport. Technical Report. 2021.
30. Bavarian Football Works. Bayern Munich Continues Injury Prevention Using Advanced Tech. 2023.
31. Kolbowicz M, Nowak M, Więckowski J. A multi-criteria system for performance assessment and support decision-making based on the example of Premier top football strikers. Phys Act Rev 2024; 12(1): 121-142. doi: 10.16926/par.2024.12.12