From High-Speed Running to Hobbling on Crutches: A Machine Learning Perspective on the Relationships Between Training Doses and Match Injury Trends

Martin Buchheit, [1 2 3 4] **Maxime Settembre,** [1] **Karim Hader,** [5] **Derek McHugh** [1]

[1]Kitman Labs, Dublin, Ireland
[2]Type 3.2 Performance
[3]HIITscience, Revelstoke, Canada
[4]INSEP, Paris, France
[5]Lille OSC, Lille, France

Soccer | **Association football** | **Extreme gradient boosting** | **Match injuries** | **Training load** | **Periodization**

Headline

In the high-stakes arena of elite football, injuries not only threaten the well-being and career longevity of players but also cast a significant shadow on teams' performances. The repercussions of injuries reverberate through both domestic and European league matches, often compromising a team's competitive edge and results (17). Beyond the pitch, the economic burden of injuries on clubs is immense, involving player treatments, lost revenues, and potential devaluation in player market worth (15).

High-speed running (HSR) occupies a central role in modern football training, crucial for enhancing performance and mitigating injury risks (5, 13). The significance of HSR and sprinting is not only in their execution but also in their strategic distribution over time in relation to technical contents and match schedules (5) - a concept known as periodization. As emphasized by Beato (4) and Dello Iacono (13), well-designed periodization guidelines are essential, delineating the volume, frequency, density, and timing of HSR and sprinting activities across the various phases of a soccer season and within the micro-cycles of each week. This periodization is what potentially equips players to withstand the rigorous demands of competitive play while avoiding injury.

In reinforcing the importance of these components, surveys by McCall (23) and Dello Iacono (13) revealed that practitioners in the field consider HSR and sprinting crucial for injury prevention strategies. This is supported by the work of Duhig et al. (14), who found that large and rapid increases in HSR distances above a player's two-year average can significantly increase the risk of hamstring injuries. Conversely, they noted that reducing HSR distances in the week leading up to a match could help reduce this risk. The interplay between the meticulous planning of HSR and its execution speaks to the delicate balance that must be maintained. While HSR is undoubtedly beneficial, it poses the question of quantification: what constitutes the optimal amount of HSR, and at what point does it become detrimental? Beato's (4) insight into the tailored periodization of HSR provides a framework for answering these questions, suggesting that it is the careful modulation of HSR—not just the raw distances run—that may hold the key to effective injury prevention.

Historical research ventures, particularly those focused on max speed exposures, have sought to pinpoint this 'optimal dose' (12, 24, 21). However, these endeavors often bear inherent limitations, constrained to specific contexts or singular populations, thus lacking a comprehensive perspective applicable across varying scenarios in football. Gualtieri et al. (16) recently advanced the discourse in this field through a comprehensive review targeting the formulation of training strategies to optimize the programming of HSR and sprinting distances in professional adult soccer. Their methodology involved examining the HSR distances traversed over the training week as a proportion of match demands. This review is particularly significant as it aligns with and extends upon previous research, which exhibits a degree of variability regarding the optimal training-to-match HSR demands ratio. For instance, Baptista (3) proposed a ratio of 0.6, suggesting a more conservative approach to HSR in training relative to match demands. Stevens et al. (28) identified a higher ratio of 1.1, indicating a nearly equivalent emphasis on HSR in both training and match play. Extending further, Martin Garcia et al. (22) presented a ratio of 1.7, while Clemente et al. (11) reported an even higher benchmark of 1.8, both advocating for a substantial volume of HSR in training exceeding that of match play. This spread of ratios reflects the ongoing discourse and diverse methodologies employed across studies in elite football, emphasizing the need for individualized approaches considering team-specific contexts and player responses. Concurrently, Silva's recent review synthesized the information contained in 16 studies to illuminate the within-microcycle distribution of HSR distance (27). However, Silva's exploration, while pioneering, lacked tangible injury data, limiting the scope and applicability of their findings.

Aim

The present study aimed to bring a fresher and deeper look at the association between HSR volumes during the training week and non-contact muscle injuries during the match ending the microcycle. Embracing a broader perspective, our present study, building upon earlier research on max speed exposures (8) and rest days (9), amalgamates data from a wide array of clubs. This expansive approach not only enhances our sample size but also enriches our analysis with diverse contexts. Distinguishing itself from previous studies, we, for the first time, integrate machine learning methods to unearth deeper patterns in football training and injury correlation. While past studies like those by López-Valenciano (19), Ayala (1), Haller (18), and Rossi (25) have employed machine learning models for injury forecasting, their relevance was limited due to two primary factors. Firstly, these studies did not boast the extensive and varied dataset we utilized. López-Valenciano (19) combined soccer and handball players, a mix that could lead to misleading conclusions due to the different training regimens and physiological characteristics of these athletes. Ayala (1) focused on just four professional soccer teams, Haller (18) on a single youth soccer team, and Rossi (25) on one professional team, thereby restricting the breadth of their findings. Secondly, previous analyses often yielded predictive insights without tangible action plans, lacking in actionable outcomes for practitioners. Furthermore, they did not delve into the dynamics of match turnaround, a critical aspect of how coaches periodize training (7,8,9). Our study provides for the first time actionable insights into the complex dynamics of training periodization and injury risk. By doing so, we aim to equip practitioners with comprehensive tools to fine-tune training programs, ultimately aiming to diminish match injuries and enhance player safety and performance. Our inclusive and technologically sophisticated approach offers a more nuanced understanding of the variegated football environments, capturing the intricacies of this sport in a unified and insightful narrative.

Methods

Study design and procedures

The overall research was based on retrospective analyses of both match injury occurrences and players' training locomotor (running) activities collected via an online database (i.e., Kitman Labs platform, Dublin, Ireland) commonly used by all the football (soccer) teams involved in the study (8,9).

Population

The elite adult football players (goalkeepers excluded) from whom data was examined belonged to 19 different teams competing in the EPL, the Italian Serie A, the French Ligue 1, the Bundesliga, the Scottish Premiership, the MLS, and the Dutch Eredivisie (from January 2018 to December 2021, with the extended covid period - from covid break (March 2020) to the end of the 2020 calendar year (next Christmas break) - being excluded from the analysis). This initial sample represented 84 team-seasons (8, 9). Then various exclusion criteria were applied at the season level:

1. Those with no or insufficient injury information (see injuries section)
2. An insufficient number of players with regular exposure throughout the season (considering that 15 is the lowest number of players that may constitute a squad)
3. Daily locomotor load not defined or no HSR and sprinting metric provided while a pitch training session or a match was registered in the team calendar.

Data extraction and anonymity

Each player and club is provided with an ID number on the platform. The researchers in charge of the analysis could only pull and analyze data associated with their IDs - no names included. Then, data was transformed and coded for injury occurrence (dates only used for assessing occurrences, such as during a match vs during training and when in relation from/to the previous match) and type (contact or non-contact injury, without any more details), to provide a final dataset. The medical staff of each team registers injury details in the platform as a part of their daily player care management, including variables such as date of injury, type and location of the injury, as well as severity (days lost). Similarly, players' match and training session participation are recorded as part of the team staff's daily monitoring. Additionally, the measures of training and competitive load are also added to the platform. The fact that all clubs used the same platform ensured the standardization and the reliability of all types of entries, from medical information to exposure measures (e.g., session duration and GPS data attached to the system calendar). We nevertheless ran a thorough data health check to ensure that all data retained for analysis met the same standard. In addition to all the steps above that guaranteed high levels of both data security and anonymity (`https://www.kitmanlabs.com/privacy-security-and-compliance/`), permission was granted by the teams for their inclusion in this research study, therefore ethics committee clearance was not required (31).

Turnarounds

A n-d turnaround was defined as a microcycle with n days between the first and second match, where n is the count of days from the first day after a match up to and including the following match day. The shortest observed turnaround was 3 days (3-d) e.g. playing a match on Sunday and again the following Wednesday, while the longest was 8 days (8-d) e.g. playing on Saturday and again the following Sunday. The longer and less common turnarounds (e.g., ≥9 days, likely including international breaks or holidays, when the training dynamics are completely different than during typical in-season turnarounds) were excluded from the analysis.

Injuries and turnaround participation

In this study we focused on non-contact match injuries as registered by the medical staff of each club, using the Orchard Sports Injury and Illness Classification System (OSIICS) offered by the online platform. While the exact diagnostic methods are impossible to describe in detail given the large variability of staff involved (i.e., 19 teams with likely more than 25 to 30 practitioners in total), the large majority of teams (if not all) at the elite football level have access to high-quality scans (i.e., Echography, Magnetic Resonance Imaging). In the literature, an injury is often defined as an occurrence sustained during either training or match-play which prevents a player from taking part in training or match-play for 1 or more days following the occurrence (2). In this study, in contrast, we wanted to focus on injuries that substantially impact training and match participation; so we only considered match injuries that caused a minimum of 3 days of training/playing interruption, i.e. ≥3-day time loss. In fact, we excluded all mild injuries (<2 days lost) because injuries in this category could conceivably not have an impact on the next game availability or training dynamic within the same turnaround. In addition, this choice has allowed us to exclude non-substantial injuries that may have resulted in a few days of unavailability due to potential training removal, as it sometimes happens in clubs (i.e., this refers to load management, when players are not injured but taken out of training for precaution – which generally allows them to train fully the next day). If the medical staff registered injuries from the start to the end of the season, we assumed that they strictly adhered to this practice throughout the whole season and that there was no missing data for this metric in this situation.

Overall, only the data of players who started and played ≥60 min in the first match of the turnaround were used for analysis, and that was considered as a 'player-turnaround'. All player turnarounds in which injuries other than non-contact time loss (≥3-d) match injuries were removed from the analysis.

Training locomotor load

Based on the above-mentioned data extract, the locomotor load metrics were used to identify the key factors that could have a relationship with injuries (4, 16):

- Accumulated HSR distance (i.e., >20 km/h) over the preceding training week
- Accumulated spriting distance (i.e., >25 km/h) over the preceding training week

Those weekly training totals were then expressed as a ratio of match demands for each individual player (16). A player's match locomotor load was assessed as the average load from the full matches he played during a given season. A full match was defined as any participation of at least 90 minutes, or >9kms when match duration was not available. When a player did not play any full match during a given season, we used the average match locomotor load from his position group during the given season. If his position group was undefined in the system, we used the average load from the whole team during the given season. Importantly, if different systems may have been used to track players' locomotor load during training and matches (e.g., GPS at training and semi-automatic video camera systems during matches), data were integrated using available equations from both the literature (6, 29) and KitmanLabs Research Initiative database (unpublished data).

Final inclusion criteria

For the purposes of this analysis, we concentrated on the most frequently occurring situations involving starter players. This decision was also influenced by the fact that monitoring substitutes' compensation is often inconsistent, especially since GPS devices are not always worn in stadiums post-match. The final dataset included 12 Teams (EPL, Championship, Bundesliga, Serie A, Ligue 1, Eredivisie, Scottish Premiership, MLS), 734 season-players (over 1-3 seasons/club) for a total of 44000 exposures (7500 matches), and 172 non-contact injuries that happened during the 2nd match of the different turnarounds examined.

Data Analysis

Descriptive analysis

For the first part of the descriptive analysis, we looked at daily locomotor loads. We examined the daily distribution of HSR and sprinting distance for the players who started and played ≥60 min during the first match of the turnaround. The daily locomotor load was then expressed as a ratio of the individual player's match demands - as defined above.

The second part of the analysis focused on the total (aggregated) turnaround training locomotor load as a function of match demands. Similarly to the first part, the players who started and played at ≥60 min during the first match of the turnaround, but also ≥60 min - or got injured - during the second match of the turnaround.

Modeling

To study the influence of high-speed and sprint running distances during the training phase on match injuries, we used eXtreme Gradient Boosting (XGBoost) for Binary Classification combined with SHAP (20) for interpretation. We took this approach as all indications from our exploratory analysis were that the relationship between injury and the explanatory variables was non-linear, and this allowed us to explore that relationship via the SHAP dependency of the variables. Those two explanatory variables were expressed as a ratio of match demands as defined in the "Training locomotor load" section.

Considering that the teams we included in this analysis have a high standard of recording their exposures, we assumed that a day without information about running distance is not missing data but rather a day where the athlete did not run over 20 or 25 km/h. We then split the data set into training and test sets in an 80:20 split, stratifying the data by teams. In addition to this, the XGBoost model structure is defined by the following hyper-parameters :

- Learning rate: 0.1
- Maximum depth of a tree: 10
- Number of trees: 100

To explain how our model works, i.e. the decisions it is making, we used SHapley Additive exPlanations (SHAP) values. SHAP does this by using fair allocation results from cooperative game theory to allocate credit for a model's output among its input features. Its calculation involves averaging the marginal contributions of each player (or feature) across all potential permutations of players (features). This involves assessing every possible combination of features and determining the impact each feature has on the model's prediction when included in these combinations. By averaging these contributions across all possible feature arrangements, it achieves a balanced and interpretable evaluation of each feature's importance in the model's prediction.

One of the fundamental properties of SHAP values is that all the input features will always sum up to the difference between baseline (expected) model output and the current model output for the prediction being explained. In the case of binary classification, the sum of the feature SHAP values equals the prediction log-odds.

As such the individual SHAP values are difficult to interpret by themselves so we made us of the feature dependence concept. As an alternative to partial dependence plots and accumulated local effects, it focuses on a given feature and shows for each data instance the relationship between the feature value (x-axis) and the corresponding SHAP value (y-axis). Based on this, we created a modified version of the SHAP dependence plot (20) (Figures 7 and 8). One dot is the SHAP value (y-axis) related to an actual observation of a given metric (x-axis) - e.g. training/game HSR distance ratio. For instance, this ratio is equal to 0.5 multiple times (Figure 7). This does not mean that all of these points get the same SHAP value as they come from different turnarounds. For each turnaround, HSR distance = 0.5 can have a different impact on the expected result depending on the other metrics' value. Local trends between two entities are shown using locally-linear estimated scatterplot smoothing (LOESS) and its 95% prediction intervals (PI). We defined four zones to summarise the impact of the feature on match injury risk:

- Increasing: both the LOESS curve and PI are above 0
- Likely increasing: the LOESS curve is above 0 but 0 is included in PI
- Likely reducing: the LOESS curve is below 0 but 0 is included in PI
- Reducing: both the LOESS curve and PI are below 0

Finally, we evaluated the probabilistic prediction of an individual match injury risk using the Brier score over the entire test set. The smaller the Brier score, the better the predictions. We compared our trained model with a naive model where all predictions are equal to the percentage of injuries (i.e. baseline) in the training set. Knowing that the model performance can vary depending on the test set, we generated 95% confidence intervals (CI) for the Brier score by bootstrapping the test set.

While we possessed data from a range of match turnarounds, spanning from 3 to 9 days, we strategically limited certain segments of the analysis, specifically those involving machine learning models, to the 6- to 8-day turnarounds. This decision was driven by two main factors. Firstly, these were the durations for which we had the most substantial sample size. Secondly, as highlighted in some of our findings (see results section), the shorter turnarounds often presented with negligible or completely absent high-speed running and sprinting distances. Such scarcity rendered the utilization of machine learning models for analysis not only impractical but also unfeasible.

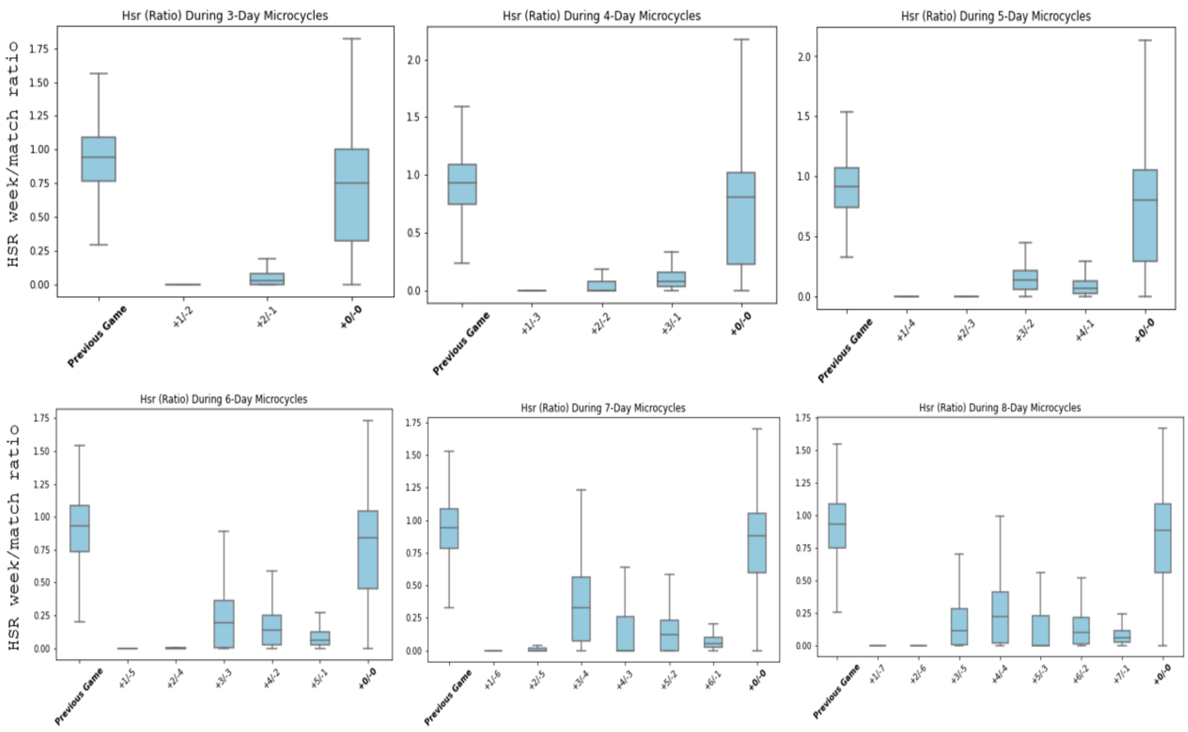

Fig. 1. High-speed running distance (>20 km/h) covered on each training day of the turnaround for each of the turnarounds examined.

Results

Figure 1 shows the daily HSR distance covered by players who started and played ≥60 min during the first match of the turnaround - for each length of the different turnarounds examined. Accumulated HSR covered during the turnaround tended to be higher in the middle of the turnarounds, especially when there were at least five days between matches (i.e., 6-d turnaround, Figure 1). It is only when the turnaround is 6-d long or more that accumulated HSR reaches at least one 1x match demands (Figure 2). However, there is a very large viability between teams when looking at the distribution of the ratio (Figure 3). When it comes to sprint running, similar trends were observed, especially in the fact that there is almost no sprint running distance at all for the 3-to-5-d turnarounds (Figure 4). The accumulated sprinting distance distribution is also similar to that described for high-speed running (Figure 2), with only turnarounds including more than 6 to 7 days between matches reaching at least a 1x match load (Figure 5). There is also a very large between-team variability with respect to how sprint distance is accumulated within each turnaround (Figure 6).

Figures 7 and 8 show the SHAP feature dependence plots for injury risk versus the accumulated training HSR and sprinting running distances during 6-to-8-d turnarounds analyzed together. The main conclusion of the modeling analysis is that there might be an association between accumulated HSR and sprinting distance during the training days over the turnarounds and match injury occurrence. More precisely, accumulated HSR distance between 0.6 and 0.9 match load, and accumulated sprinting distance between 0.6 and 1.1 match load, tended to be associated with a lower injury risk.

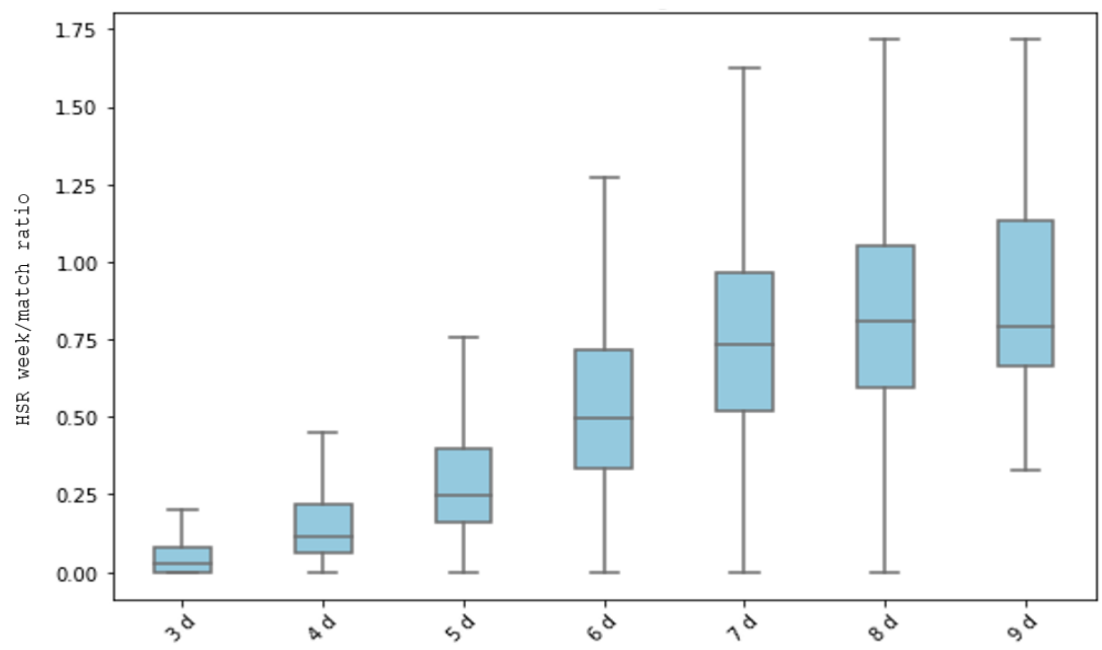

Fig. 2. Cumulated high-speed running distance (>20 km/h, expressed as a ratio of match demands) during training for each turnaround (match excluded).

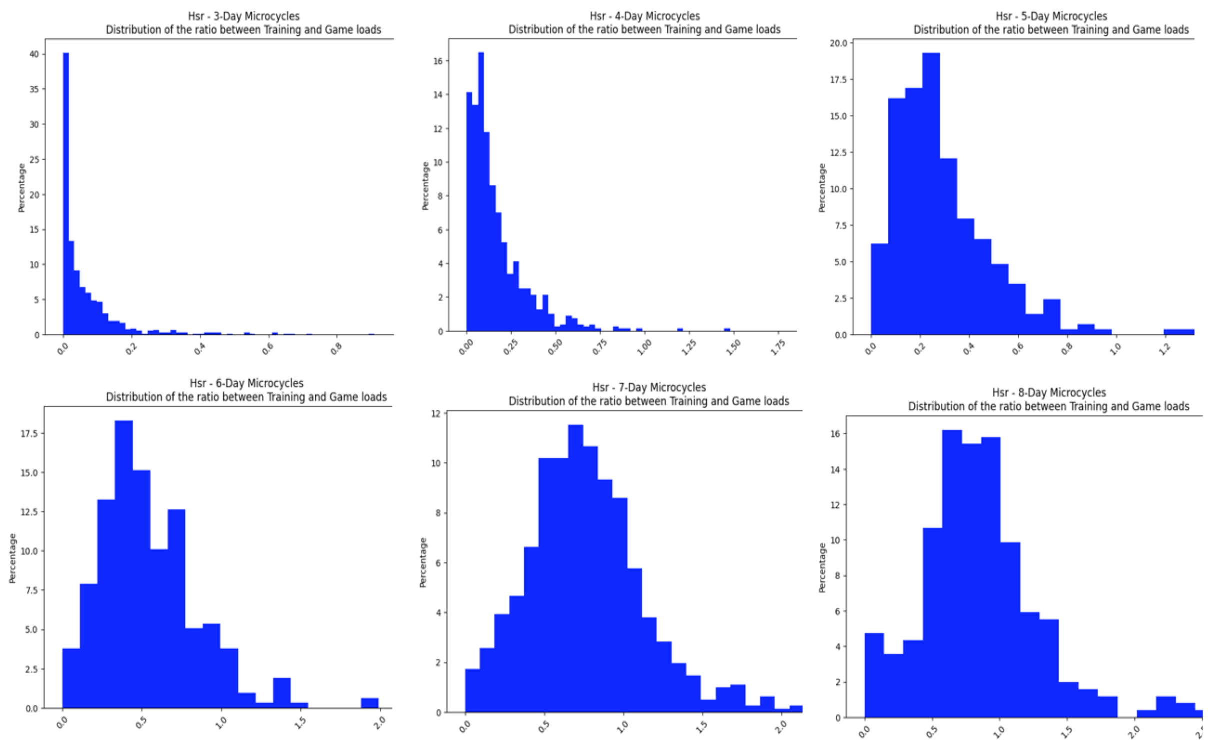

Fig. 3. Distribution of the cumulated high-speed running distance (>20 km/h) during training for each turnaround in relation to match demands (i.e., expressed as a ratio).

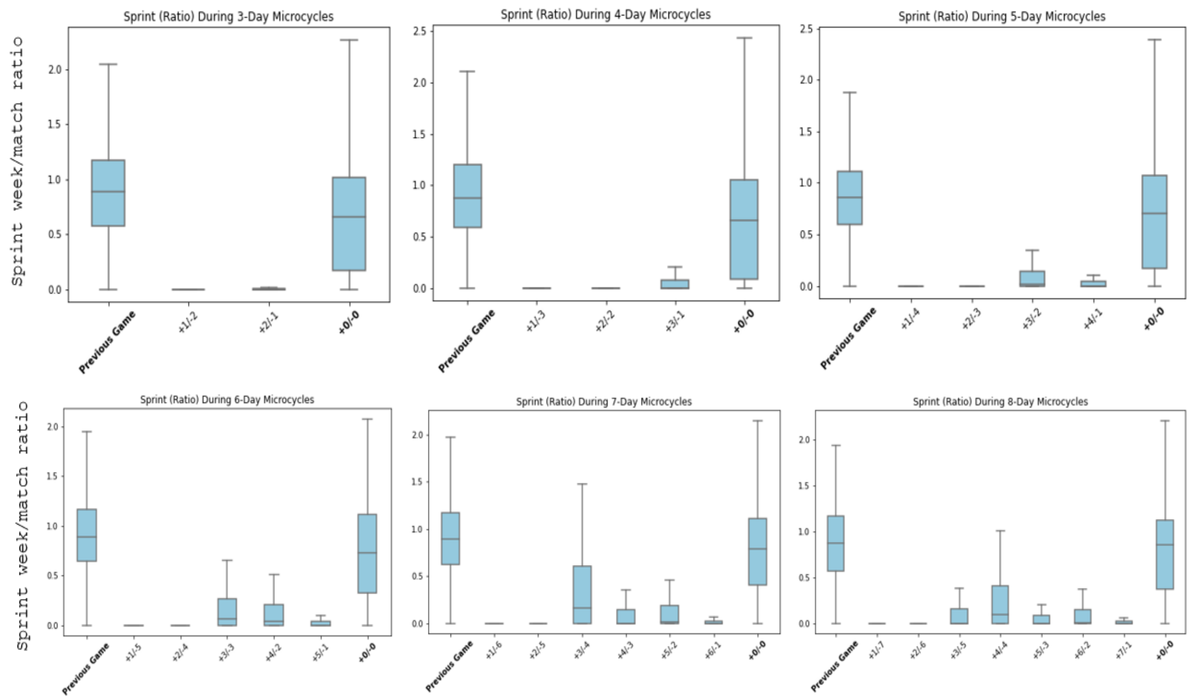

Fig. 4. Sprint running distance (>25 km/h) covered during training for each day of the turnaround for each of the turnaround examined.

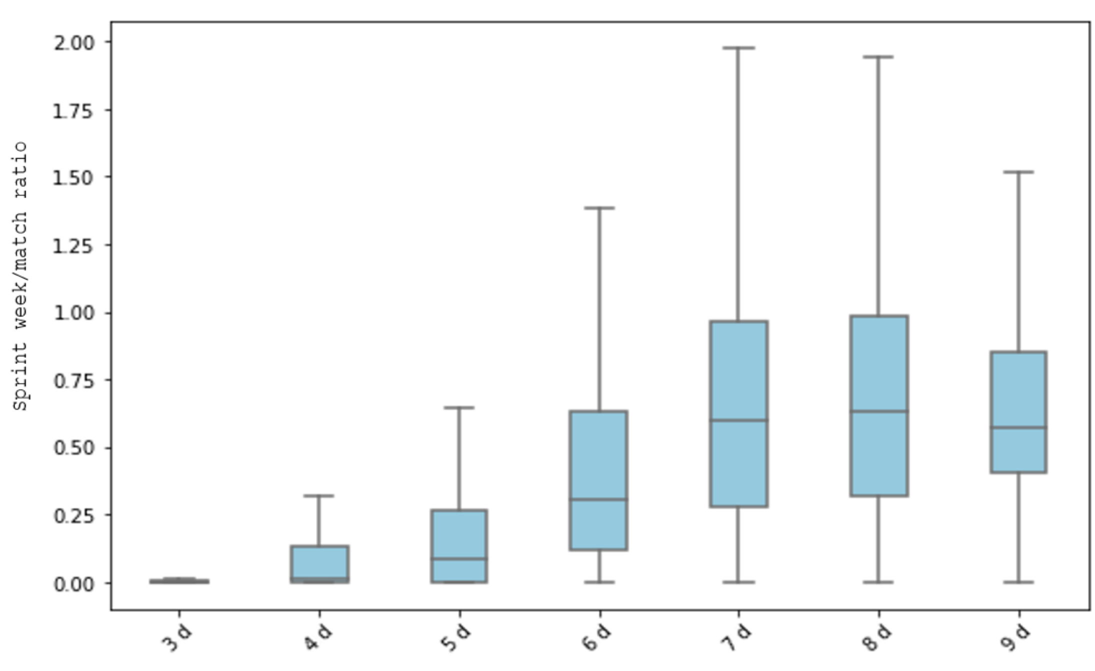

Fig. 5. Cumulated sprint running distance (>25 km/h, expressed as a ratio of match demands) during training for each turnaround.

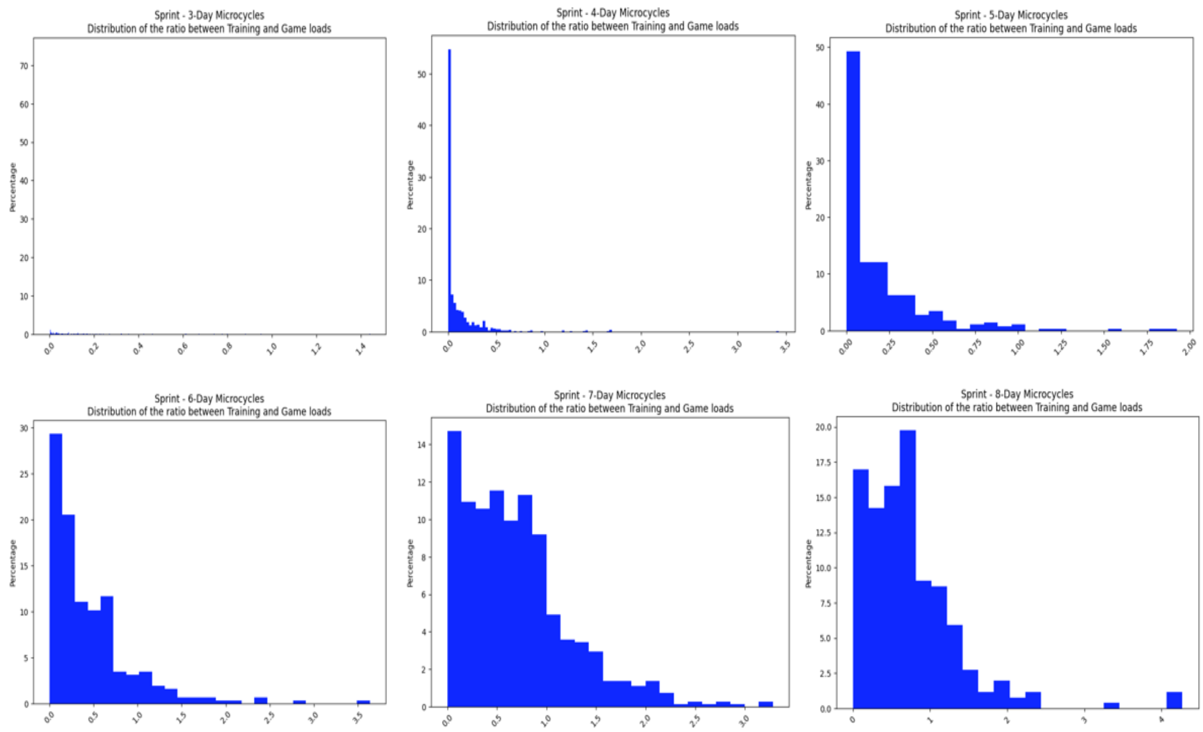

Fig. 6. **Distribution of the cumulated sprint running distance (>25 km/h) during training for each turnaround in relation to match demands (i.e., expressed as a ratio).**

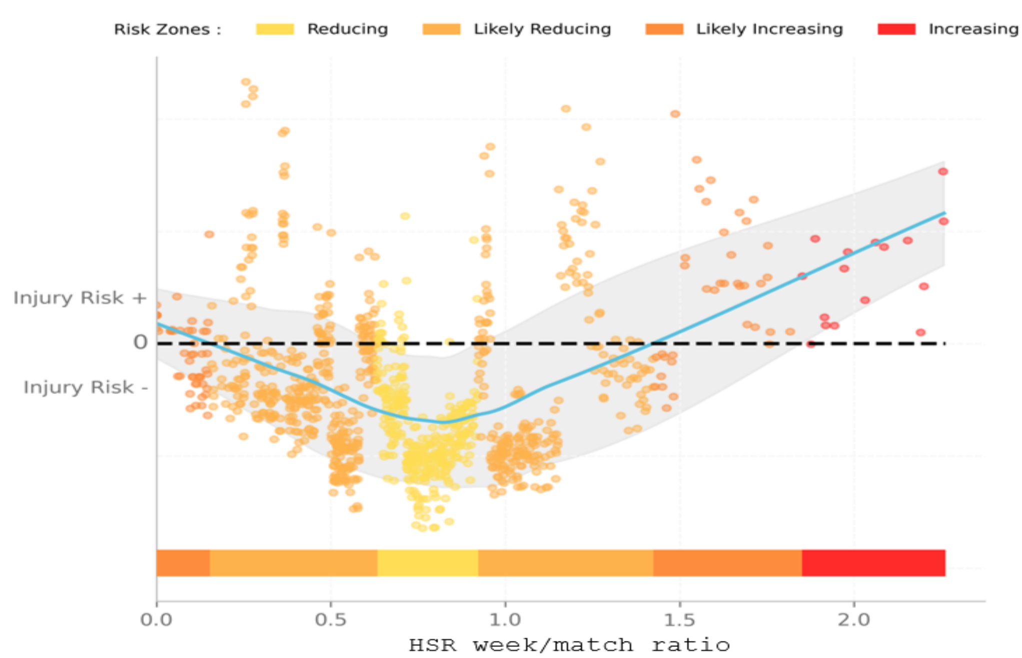

Fig. 7. **SHAP feature dependence plots for match injury risk vs. cumulated high-speed running distance (>20 km/h) during training over 6- to 8-day turnarounds (expressed as a ratio of match demands). Injury (+/-) is quantified as the magnitude of the SHAP contribution.**

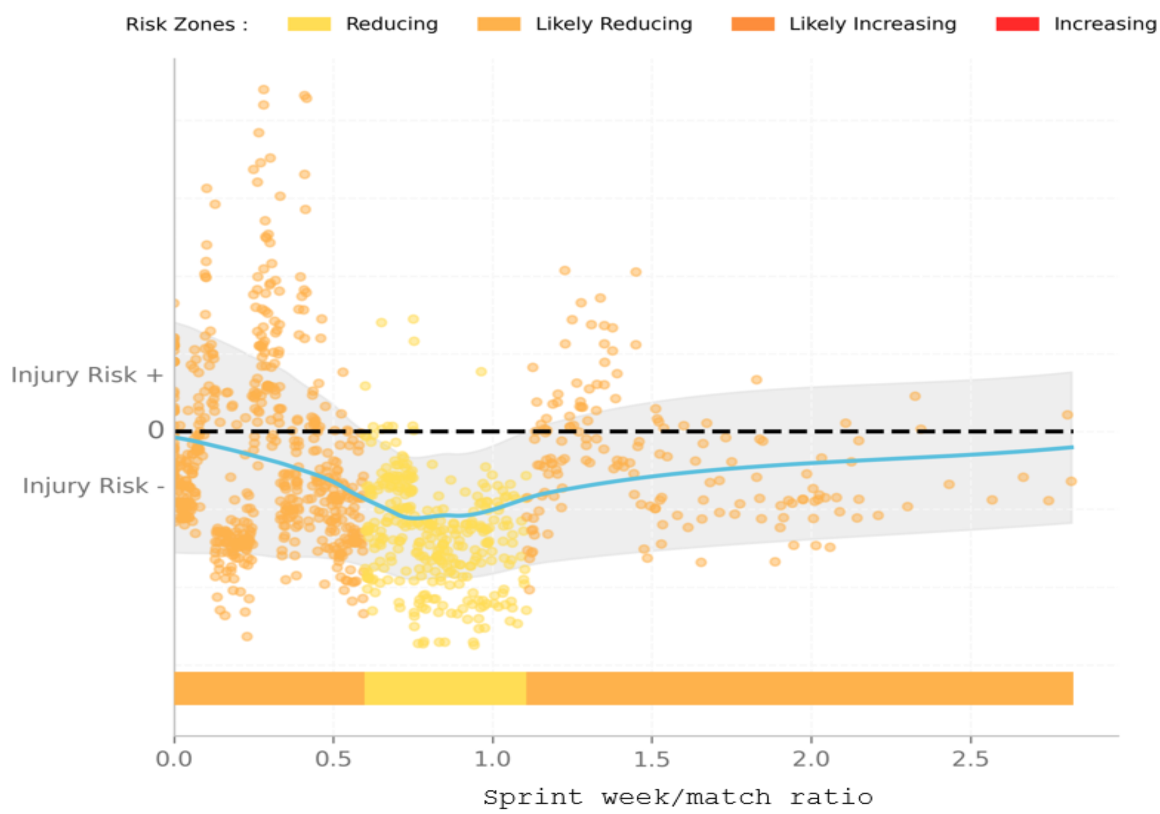

Fig. 8. SHAP dependency plots for match injury risk vs. cumulated sprint running distance (>25 km/h) during training over 6- to 8-day turnarounds (expressed as a ratio of match demands). Injury (+/-) is quantified as the magnitude of the SHAP contribution.

Discussion

The implications of injuries in elite football extend beyond the immediate physical toll on players. As emphasized in our introduction, these injuries can substantially influence teams' performances in both domestic and European league matches (17). Furthermore, the resultant economic burden on clubs from treatment, lost revenues, and potential devaluation in player market value is staggering (15). Therefore, finding methods to reduce injury rates, particularly through understanding training loads like HSR and sprinting, becomes paramount (23).

Our study, leveraging a comprehensive dataset from multiple elite European and MLS football clubs and machine learning techniques, elucidated intricate patterns in the relationship between match turnaround durations, HSR, sprinting distances, and their potential association with match injuries. Our study distinguishes itself in the realm of football training and injury analysis by utilizing a comprehensive dataset from multiple clubs, unlike previous narrower-focused studies. While earlier research in adult pro football (1, 25) also used machine learning for injury prediction, our approach provides actionable insights into training periodization, particularly in match turnaround dynamics. This not only enhances the practical application of our findings but also offers a nuanced understanding crucial for reducing injury rates and improving player safety and performance.

A noteworthy finding from our research highlighted a pronounced peak in accumulated HSR during the middle of the turnaround, particularly evident in situations with turnarounds of six days or more (Figures 1 and 4). This pattern mirrors the observations made in Silva's comprehensive review (27), which aggregated data from 16 studies, revealing a peak in HSR demands typically around 4 and 3 days before the match (MD-4 and MD-3). However, it's crucial to note that Silva's findings did not incorporate an injury context.

Our observations also align with a survey conducted by Buchheit (7), wherein 100 elite football practitioners collectively conveyed the necessity of a minimum of five training days between matches for the effective incorporation of intense sessions, including HSR. The challenge posed by shorter intervals between matches is the limited recovery timeframe post-match, compounded by the subsequent reduced recovery period before the next intense training session. This sentiment was further echoed by the feedback from 99 practitioners surveyed by Dello Iacono (13). Their preferred training schedule for starting players typically spanned from at least 48 hours post the previous match (MD + 2) to no later than 48 hours before the upcoming match (MD - 2). However, diving deeper into our findings, particularly considering the classic 7-day turnaround, it's intriguing to observe that the highest HSR and sprinting demands materialized on MD-4 rather than MD-3. Conventionally, MD-4 is reserved for strength-focused sessions, characterized by smaller-sized games, whereas MD-3 is generally earmarked for endurance training with larger spaces utilized. This observed deviation is intriguing. While the precise reasons remain elusive, it's conceivable that they are rooted in the unique periodization contexts of individual teams, influenced by factors like designated rest days during the turnaround.

Our study's distinct emphasis, when compared to more narrowly-scoped prior research, hinges on the notable between-team variability in the training HSR and sprinting volumes of starter players, and how these might relate to match injuries. This variability is discernible at the daily programming level, as demonstrated by the fluctuations observed on specific days (Figures 1 and 4), and also in the context of match demands, illustrated by variances across different turnarounds (Figures 2, 3, 5, and 6). These differences underscore that universally applicable strategies remain elusive, a sentiment echoed by Gualtieri (16). In the present data set, despite some extreme values close to 0 or >2, the large majority of the teams showed training-to-match HSR demand ratios between 0.5 and 1 (Figure 3). This is similar to the range of values reported previously (e.g., Baptista (3): 0.6; Stevens (28): 1.1; Martin Garcia (22): 1.7 and Clemente (11): 1.8). It's this very variability and its implications that we've delved into, aiming to understand how different training-to-match HSR demand ratios correlate with varying levels of injury risk.

Drawing parallels to the surveys by McCall (23), Dello Iacono (13) and the study by Duhig et al. (14), it's evident that while HSR and sprinting are crucial components of player training and performance, their dosage and programming require careful consideration (5). Our findings suggest for the first time that accumulated HSR distances between 0.6 and 0.9 match load (Figure 7), and sprinting distances between 0.6 and 1.1 match load (Figure 8), might be associated with decreased injury occurrences. This suggests there may be an optimal range within which players could benefit from improved performance while mitigating injury risks.

It's essential, however, to approach these results with measured optimism. While our machine learning models provide deeper insights than traditional methods, we must remember that correlation does not equate to causation (30). The performance of the model compares to that of a naive approach so there is more work to be done to isolate any effect. Yet, given the comprehensive nature of our data and the advanced methods employed, we believe our findings can be pivotal for practitioners. Furthermore, while our study offers insights into the effects of changes in the ratio between training and match demands on match injuries (Figures 7 and 8), future research should delve deeper. The results presented here indicate that a more rigorous experimental set-up, in as controlled an environment as possible in elite sport, would be fruitful and we strongly encourage any proposals in this area. Exploring how the distribution of locomotor load throughout the microcycle (Figures 1 and 4) might influence these injury outcomes can provide a more holistic understanding of injury prevention in the context of varied training dynamics. For example, whether it may be better to program 1 match load of HSR and sprinting distance on a single day vs. spread over 2 or 3 days, is still unknown, and did not reach a consensus among the practitioners surveyed (7, 13).

Conclusions

In our pioneering study using a comprehensive dataset from elite European football, we delved into the intricate relationship between match turnaround durations, distances covered in high-speed running (HSR) and sprinting, and their association with match injuries. This innovative analysis, the first of its kind to amalgamate data from such a vast number of teams, employed machine learning models, marking a novel approach to link HSR and sprinting loads with injury risks. Our models pinpointed specific ranges for starter players: HSR distances of 0.6 to 0.9 match load and sprinting distances of 0.6 to 1.1 match load, which exhibited a correlation with diminished injury occurrences. Beyond the weekly totals, further studies should also examine the effect of the distribution of this locomotor load across the microcycle on the present results.

References

1. Ayala F, López-Valenciano A, Gámez Martín JA, De Ste Croix M, Vera-Garcia FJ, García-Vaquero MDP, Ruiz-Pérez I, Myer GD. A Preventive Model for Hamstring Injuries in Professional Soccer: Learning Algorithms. Int J Sports Med. 2019 May;40(5):344-353. doi: 10.1055/a-0826-1955. Epub 2019 Mar 14.

2. Bahr R, Clarsen B, Derman W, Dvorak J, Emery CA, Finch CF, Hägglund M, Junge A, Kemp S, Khan KM, Marshall SW, Meeuwisse W, Mountjoy M, Orchard JW, Pluim B, Quarrie KL, Reider B, Schwellnus M, Soligard T, Stokes KA, Timpka T, Verhagen E, Bindra A, Budgett R, Engebretsen L, Erdener U, Chamari K. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sport 2020 (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). Br J Sports Med. 2020;54(7):372-389.

3. Baptista I, Johansen D, Seabra A, Pettersen SA. Position specific player load during match-play in a professional football club. PLoS One. 2018 May 24;13(5):e0198115. doi: 10.1371/journal.pone.0198115. eCollection 2018.

4. Beato M, Drust B, Iacono AD. Implementing High-speed Running and Sprinting Training in Professional Soccer. Int J Sports Med. 2021 Apr;42(4):295-299. doi: 10.1055/a-1302-7968. Epub 2020 Dec 8.

5. Buchheit M. Programming high-speed running and mechanical work in relation to technical contents and match schedule in professional soccer. Sport Performance & Science Reports, 2019, July, #69, V1.

6. Buchheit M, Allen A, Poon TK, Modonutti M, Gregson W, Di Salvo V. Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. J Sports Sci. 2014 Dec;32(20):1844-1857. Doi: 10.1080/02640414.2014.942687. Epub 2014 Aug 5.

7. Buchheit M, Sandua M, Berndsen J, Shelton, Smith S, Norman D, McHugh D and Hader K. Loading patterns and programming practices in elite football: insights from 100 elite practitioners. Sport Perf & Sci Research. 2021;(153):v1.

8. Buchheit M, Settembre M, Hader K, McHugh D. Exposures to near-to-maximal speed running bouts during different turnarounds in elite football: association with match hamstring injuries. Biol Sport. 2023 Oct;40(4):1057-1067. doi: 10.5114/biolsport.2023.125595. Epub 2023 Mar 6.

9. Buchheit M, Settembre M, Hader K, McHugh D. Planning the Microcycle in Elite Football: To Rest or Not to Rest? Int J Sports Physiol Perform. 2023 Jan 3;18(3):293-299. doi: 10.1123/ijspp.2022-0146. Print 2023 Mar 1.

10. 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 Med Open. 2019 Jul 3;5(1):28. doi: 10.1186/s40798-019-0202-3.

11. Clemente FM, Rabbani A, Conte D, Castillo D, Afonso J, Truman Clark CC, Nikolaidis PT, Rosemann T, Knechtle B. Training/Match External Load Ratios in Professional Soccer Players: A Full-Season Study. Int J Environ Res Public Health. 2019 Aug 23;16(17):3057. doi: 10.3390/ijerph16173057.

12. Colby MJ, Dawson B, Peeling P, Heasman J, Rogalski B, Drew MK, Stares J. Improvement of Prediction of Noncontact Injury in Elite Australian Footballers With Repeated Exposure to Established High-Risk Workload Scenarios. Int J Sports Physiol Perform. 2018;13(9):1130-1135.

13. Dello Iacono A, Beato M, Unnithan VB, Shushan T. Programming High-Speed and Sprint Running Exposure in Football: Beliefs and Practices of More Than 100 Practitioners Worldwide. Int J Sports Physiol Perform. 2023 Apr 28;18(7):742-757. doi: 10.1123/ijspp.2023-0013. Print 2023 Jul 1.

14. Duhig S, Shield AJ, Opar D, Gabbett TJ, Ferguson C, Williams M. Effect of high-speed running on hamstring strain injury risk. Br J Sports Med. 2016 Dec;50(24):1536-1540. doi: 10.1136/bjsports-2015-095679. Epub 2016 Jun 10.

15. Eliakim E, Morgulev E, Lidor R, Meckel Y. Estimation of injury costs: financial damage of English Premier League teams' underachievement due to injuries. BMJ Open Sport Exerc Med. 2020 May 20;6(1):e000675. doi: 10.1136/bmjsem-2019-000675. eCollection 2020.

16. Gualtieri A, Rampinini E, Dello Iacono A, Beato M. High-speed running and sprinting in professional adult soccer: Current thresholds definition, match demands and training strategies. A systematic review. Front Sports Act Living. 2023 Feb 13;5:1116293. doi: 10.3389/fspor.2023.1116293. eCollection 2023.

17. Hägglund M, Waldén M, Magnusson H, Kristenson K, Bengtsson H, Ekstrand J. Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. Br J Sports Med. 2013 Aug;47(12):738-42. doi: 10.1136/bjsports-2013-092215. Epub 2013 May 3.

18. Haller N, Kranzinger S, Kranzinger C, Blumkaitis JC, Strepp T, Simon P, Tomaskovic A, O'Brien J, Düring M, Stöggl T. Predicting Injury and Illness with Machine Learning in Elite Youth Soccer: A Comprehensive Monitoring Approach over 3 Months. J Sports Sci Med. 2023 Sep 1;22(3):476-487. doi: 10.52082/jssm.2023.476. eCollection 2023 Sep.

19. López-Valenciano A, Ayala F, Puerta JM, DE Ste Croix MBA, Vera-Garcia FJ, Hernández-Sánchez S, Ruiz-Pérez I, Myer GD. A Preventive Model for Muscle Injuries: A Novel Approach based on Learning Algorithms. Med Sci Sports Exerc. 2018 May;50(5):915-927. doi: 10.1249/MSS.0000000000001535.

20. Lundberg S., Lee S-I., (2017), A Unified Approach to Interpreting Model Predictions. arXiv.1705.07874

21. Malone S, Roe M, Doran DA, et al. High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football. J Sci Med Sport. 2017;20:250-4

22. Martín-García A, Gómez Díaz A, Bradley PS, Morera F, Casamichana D. Quantification of a Professional Football Team's External Load Using a Microcycle Structure. J Strength Cond Res. 2018 Dec;32(12):3511-3518. doi: 10.1519/JSC.0000000000002816.

23. McCall A, Pruna R, Van der Horst N, Dupont G, Buchheit M, Coutts AJ, Impellizzeri FM, Fanchini M; EFP-Group. Exercise-Based Strategies to Prevent Muscle Injury in Male Elite Footballers: An Expert-Led Delphi Survey of 21 Practitioners Belonging to 18 Teams from the Big-5 European Leagues. Sports Med. 2020 Sep;50(9):1667-1681. doi: 10.1007/s40279-020-01315-7.

24. O'Connor F, Thornton HR, Ritchie D, et al. Greater association of relative thresholds than absolute thresholds with non-contact lower-body injury in professional Australian rules footballers: Implications for sprint monitoring. Int J Sports Physiol Perform. 2019;5(2):204-212.

25. Rossi A, Pappalardo L, Cintia P, Iaia FM, Fernàndez J, Medina D. Effective injury forecasting in soccer with GPS training data and machine learning. PLoS One. 2018 Jul 25;13(7):e0201264. doi: 10.1371/journal.pone.0201264. eCollection 2018.

26. Settembre M, Buchheit M, Hader K, Hamill R, Tarascon A, Verheijen R, McHugh D. Factors associated with match outcomes in elite European football - insights from machine learning models. Journal of Sports Analytics. In press

27. Silva H, Nakamura F, Castellano J, Marcelino R. Training Load Within a Soccer Microcycle Week—A Systematic Review. Strength and Conditioning Journal 45(5):p 568-577, October 2023. | DOI: 10.1519/SSC.0000000000000765

28. Stevens TGA, de Ruiter CJ, Twisk JWR, Savelsbergh GJP, Beek PJ. Quantification of in-season training load relative to match load in professional Dutch Eredivisie football players. Sci Med Football. 2017 1(2):117–25. doi: 10.1080/24733938.2017.1282163

29. Taberner M, Allen T, O'Keefe J, Richter C, Cohen D, Harper D, Buchheit M. Interchangeability of optical tracking technologies: potential overestimation of the sprint running load demands in the English Premier League. Sci Med Football. 2022 Aug 8:1-10. doi: 10.1080/24733938.2022.2107699.

30. Wilson EB. Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association. 1927 22(158): 209–212.

31. Winter EM, Maughan RJ. Requirements for ethics approvals. J Sports Sci. 2009;27(10):985.