OPEN

Predicting Severity of Head Collision Events in Elite Soccer Using Preinjury Data: A Machine Learning Approach

Gabriel Tarzi, BSc,* Christopher Tarzi, BHSc,* Ashirbani Saha, PhD,*† and Michael D. Cusimano, MD, PhD*‡

Abstract

Objective: To develop machine learning (ML) models that predict severity of head collision events (HCEs) based on preinjury variables and to investigate which variables are important to predicting severity. **Design:** Data on HCEs were collected with respect to severity and 23 preinjury variables to create 2 datasets, a male dataset using men's tournaments and mixed dataset using men's and women's tournaments, to perform ML analysis. Machine learning analysis used a random forest classifier based on preinjury variables to predict HCE severity. **Setting:** Four elite international soccer tournaments. **Participants:** Elite athletes participating in analyzed tournaments. **Independent Variables:** The 23 preinjury variables collected for each HCE. **Main Outcome Measures:** Predictive ability of the ML models and association of important variables. **Results:** The ML models had an average area under the receiver operating characteristic curve for predicting HCE severity of 0.73 and 0.70 for the male and mixed datasets, respectively. The most important variables for prediction were the mechanism of injury and the event before injury. In the male dataset, the mechanisms "head-to-head" and "knee-to-head" were together significantly associated ($P = 0.0244$) with severity; they were not significant in the mixed dataset ($P = 0.1113$). In both datasets, the events "corner kicks" and "throw-ins" were together significantly associated with severity (male, $P = 0.0001$; mixed, $P = 0.0004$). **Conclusions:** ML models accurately predicted the severity of HCE. The mechanism and event preceding injury were most important for predicting severity of HCEs. These findings support the use of ML to inform preventative measures that will mitigate the impact of these preinjury factors on player health.

Key Words: traumatic brain injury, machine learning, soccer, injury prevention

(*Clin J Sport Med* 2023;33:165–171)

INTRODUCTION

Increased research on sport-related concussions (SRCs), a type of traumatic brain injury (TBI), has resulted in greater public attention toward concussions in recent years. Sport-related concussions are a significant contributor to annual TBI incidence and warrant heightened intervention.[1] Returning to play while still experiencing symptoms of a concussion increases the risk for long-term brain-health consequences.[2] Immediate removal from play and formal evaluation by medical professionals of athletes suspected of sustaining a concussion is critical to avoiding further negative brain-health consequences.[2] Recognizing and adequately assessing players involved in head collision events (HCEs) is essential to ensuring athletes receive proper care and reduce the burden of concussion in sport.

Association football, commonly known as football or soccer, which is governed by the Fédération Internationale de Football Association (FIFA), has some of the highest rates of concussions in sport.[3,4] This concern is magnified by soccer's worldwide popularity and research, which found that major international tournaments, such as the World Cup (WC), do not follow necessary assessment protocol for HCEs.[5–8] Determining the severity of HCEs can guide team medical personnel and officials to recognize potential concussive events and assure appropriate assessment, care, and treatment is provided. At the same time, identification of potential markers of severity can help guide real-time decision making on the soccer pitch as to which players and what sorts of events require more in-depth clinical assessments.

Data-driven techniques, such as machine learning (ML), are often used in the identification of various contextual markers or precursors of events in various disciplines.[9,10] In team

Submitted for publication February 21, 2022; accepted September 15, 2022.

*From the *Injury Prevention Research Office, Division of Neurosurgery, St. Michael's Hospital, Toronto, ON, Canada; †Department of Oncology, McMaster University, Hamilton, ON, Canada; and ‡Department of Surgery, University of Toronto, Toronto, ON, Canada.*

Supported by the Canadian Institutes of Health Research Strategic Team Grant in Applied Injury Research #TIR-103946 and the Ontario Neurotrauma Foundation.

The authors report no conflicts of interest.

Research concept and design by M. D. Cusimano; data collection and analysis by C. Tarzi, G. Tarzi and A. Saha; Machine learning analysis by A. Saha; manuscript writing by C. Tarzi, A. Saha, G. Tarzi and M. D. Cusimano; manuscript reviewed and approved by C. Tarzi, G. Tarzi, A. Saha, and M. D. Cusimano.

This study was exempt from ethical approval by the Research Ethics Board at St. Michael's Hospital.

Data are available upon reasonable request.

Code is available upon reasonable request.

Corresponding Author: Michael D. Cusimano, MD, MHPE, FRCSC, PhD, FACS, Division of Neurosurgery, Li Ka Shing Knowledge Institute, Unity Health, St. Michael's Hospital, Professor of Neurosurgery, Education and Public Health, University of Toronto, 30 Bond St, Toronto, ON, Canada M5B 1W8 (injuryprevention@smh.ca).

http://dx.doi.org/10.1097/JSM.0000000000001087

sports, ML is actively being used in the assessment of injury risk for athletes.[11] Injury risk prediction, particularly in soccer, has been performed using various types of data such as wearable sensors, psychosocial stress, and training load, but with no particular focus to head injuries or concussions.[12–17] Studies that have applied ML to investigate injury prediction in sports predominantly addressed male athletes, in mostly small cohorts (<100), and tended not to focus on professional athletes, who have significantly higher rates of injuries, particularly during matches compared with training.[18] This study uses data on more than 200 HCEs from male and female elite soccer tournaments to build an ML model that predicts the severity of HCEs based on preinjury variables. This novel approach to studying HCEs will not only inform decisions on the soccer pitch, but also demonstrate the capabilities of ML to be applied to other sports.

METHODS

HCE Identification

Data on HCEs were previously collected by 12 trained independent reviewers analyzing 4 major international soccer tournaments, the 2014 FIFA WC, 2016 UEFA European Championship, 2018 FIFA WC, and 2019 FIFA Women's WC.[5–8] Reviewers conducted standardized training and only proceeded to match analysis if the Cohen's kappa κ was greater than 0.85 for HCE agreement. Two reviewers independently analyzed match footage for all 231 games included in this study.

Independent reviewers identified HCEs as incidents in which a player suffered a direct head contact and was unable to resume play within 5 seconds. The term HCE is used to encompass a wide range of head collisions that merit medical assessment and can potentially result in a concussion. Events that were intentional (eg intentionally heading the ball), showed clear embellishment, lacked conclusive video evidence, or minor head contact (eg fingers lightly brushing players head) were excluded from the study. For each HCE identified, information on the players involved (eg height, weight, sex, position), preinjury situational factors (eg location of play on the pitch, game circumstances), medical assessment, and signs of concussion were collected by independent reviewers (Table 1).

Signs of potential concussion were defined by visible signs that are associated with a concussion diagnosis and were confirmed by video analysis, these include: clutching of the head, slowness getting up, disorientation, disequilibrium, loss of consciousness, and seizures. Severity is coded as a binary variable (severe/non-severe). A singular HCE is considered severe if the player involved exhibited more than 2 signs of concussion. The research ethics board at St. Michael's Hospital waived the need for ethics approval. All data were collected from publicly available information and was conducted in accordance with the relevant guidelines.

Machine Learning and Data Analysis

Dataset Division

To perform ML, 2 datasets were developed. The first dataset resulted from combining data from the male tournaments together, named the "male dataset." The second dataset was a combination of all male and female tournament data together and hence called the "mixed dataset." The specifics of analysis for each of these datasets is described below.

Male Dataset

In this dataset, events with missing data were removed from all before analyzing the data. A random forest classification schema was used to assess the ability of the 23 preinjury variables to predict the postinjury severity.[19] A nested cross-validation (CV) scheme was used to create and find the best model for prediction. We ran 100 iterations of a 4-fold CV (outer CV) on the dataset and assessed the CV performance of the classifier. Within each outer CV fold, we ran a repeated 3-fold CV (inner CV) 5 times and found the best model for prediction using random search on the hyperparameter space. This model was used to assess the prediction performance in the area-under-receiver-operating characteristics curve (AUC) in each outer fold. The average AUC (over all outer folds) was noted. The average AUCs for 100 iterations of the outer CV loop was tabulated and mean and SD of the average AUCs were calculated to estimate the performance of the classification procedure. The outer CV partitions were made by function createFolds from the package Caret in R software.[20] The average ranking of the variables in the mean decrease in Gini index was noted to identify the important markers of predictors of severity. Higher ranking of the mean decrease in Gini index of a variable indicates increased importance as a predictor. The method of random forest was chosen because of the prominent feature of the technique to rank importance of the variables for variable selection, and it is experimentally validated to be a competitive method for variable selection.[21] Moreover, our data have imbalance (17% severe incidents or minority classes for the mixed dataset) and random forest is able to perform well with this amount of imbalance without any preprocessing measures.[22] Thereafter, Fisher exact test were conducted to analyze the association of these markers with severity. All analyses were conducted in R software.[20]

Mixed Dataset

Only variables that were common between the male and female, and that had no more than 10 missing values were included. This was performed to retain most of the 23 variables that we had. Data on experience and weight for player 1 and player 2 were not available for the 2019 Women's WC and were excluded from analysis in the mixed dataset. In addition, a new variable "Sex," coded as male/female depending on the tournament of interest, was added. Then HCE severity prediction from 20 predictors was conducted in the same manner as described in the section above.

RESULTS

HCE Identification

A total of 263 HCEs were identified and coded across 231 matches in the 4 elite soccer tournaments analyzed. Across all tournaments, 45 HCEs were severe and 218 were nonsevere. A breakdown of the classification of all HCEs is presented in Table 2.

TABLE 1. List of Preinjury Variables Collected and Used in the Machine Learning Analysis, Along With the Corresponding Definitions

Variable	Definition
Player factors	
Experience of player 1*	Number of years that player 1 has been playing soccer professionally
Experience of player 2*	Number of years that player 2 has been playing soccer professionally
Age of player 1	Age of player 1 in years
Age of player 2	Age of player 2 in years
Height of player 1	Height of player 1 in cm
Height of player 2	Height of player 2 in cm
Weight of player 1*	Weight of player 1 in kg
Weight of player 2*	Weight of player 2 in kg
Position of player 1	Position on soccer field of player 1
Position of player 2	Position on soccer field of player 2
Player sex†	Sex of the player (applicable for mixed dataset only)
On-field factors	
Number of players involved	Number of players involved in injury
Number of players injured	Number of players injured (HCE)
Direction	Direction player 1 was playing in before injury
Action player 1	Action of player 1 before injury
Action player 2	Action of player 2 before injury
Event prior to injury	Type of play occurring just before injury
Mechanism of injury for player 1	Mechanism of contact for player 2 for injury 1
Mechanism of injury for player 2	Mechanism of contact for player 1 for injury 1
Location of impact for player 1	Impact location of injury on head of player 1
Location of impact for player 2	Impact location of injury on head of player 2
Area	Area on the soccer field where injury occurred
Game factors	
Game score	Whether player 1's team was winning, losing, or tied at the time of HCE
Field time	Time in match the HCE took place

** Variables not included in mixed dataset.*
† Variable not included in male dataset.

Male Dataset

For the male dataset, a total of 215 HCEs were found in 179 matches after removing any missing data from the 23 variables used as predictors. The performance of the ML classification procedure was evaluated using the mean of the average AUC in the outer CV loops, which was 0.73 (SD = 0.03). In each outer CV iteration, the preinjury variables were ranked in order of importance for classification, in mean decrease in Gini index. The average importance ranking of the 23 preinjury variables is shown in Figure 1.

Preinjury event and injury mechanism were the top 2 important variables found from our classification scheme through mean decrease in Gini index and both ranked significantly more important than the third variable. The frequency and proportion of different categories of preinjury event and injury mechanism is provided in Table 2. In a subanalysis, the preinjury events "corner kick" and "throw-in" together have a significant association with severity compared with other events (Fisher exact test, *P*-value = 0.0001122). Either of these 2 preinjury events were present in 12.5% (27/215) of HCEs and in 33.3% (13/39) of all severe HCEs. In 48.1% (13/27) of HCEs that had "Corner kick" or "throw-in," severe injury occurred. In comparison, 13.8% (26/188) of the remaining HCEs were severe. The injury mechanisms "head-to-head" and "knee-to-head" together have a significant association with severity compared with other mechanisms (Fisher exact test, *P*-value = 0.0244). Either of these 2 mechanisms were present in 19.5% (42/215) of HCEs and in 33.3% (13/39) of the severe HCEs. In 30.9% (13/42) HCEs that had "head-to-head" and "knee-to-head," severe injury occurred. In comparison, 15% (26/173) of the remaining HCEs were severe.

Mixed Dataset

For the mixed dataset, a total of 263 events were found in 231 matches with 20 predictors. The mean of the average AUC values was 0.70 (SD = 0.02). The average ranking of the 20 preinjury variables used to predict severity, in mean decrease in Gini index, is shown in Figure 2. Again, through mean decrease in Gini index analysis, injury mechanism and preinjury events were the top 2 important variables found from our classification scheme in the mixed dataset. Corner kick and throw-in together remained significantly associated

TABLE 2. Breakdown of the Characteristics of Severe and Non-Severe HCEs Collected Across all 4 Tournaments

	Severe HCE (n = 45)	Nonsevere HCE (n = 218)
Men	39 (86.7%)	176 (80.7%)
Women	6 (13.3%)	42 (19.3%)
Mechanism of injury		
Head to head	12 (26.7%)	39 (17.9%)
Hand or first to head	4 (8.9%)	43 (19.7%)
Shoulder to head	3 (6.7%)	11 (5.0%)
Elbow to head	8 (17.8%)	74 (33.9%)
Foot to head	2 (4.4%)	11 (5.0%)
Leg or hip to head	4 (8.9%)	12 (5.5%)
Other body part to head	1 (2.2%)	4 (1.8%)
Head to field	2 (4.4%)	7 (3.2%)
Hit by ball	4 (8.9%)	11 (5.0%)
Knee to head	2 (4.4%)	4 (1.8%)
Other	1 (2.2%)	1 (0.5%)
Not visible in footage	2 (4.4%)	1 (0.5%)
Preinjury event		
Corner kick	6 (13.3%)	11 (5.0%)
Goal kick	4 (8.9%)	17 (7.8%)
Short pass	14 (31.1%)	55 (25.2%)
Long pass	6 (13.3%)	50 (22.9%)
Throw-in	7 (15.6%)	7 (3.2%)
Dribbling	2 (4.4%)	47 (21.6%)
Penalty kick	0 (0%)	3 (1.4%)
Other	6 (13.3%)	28 (12.8%)

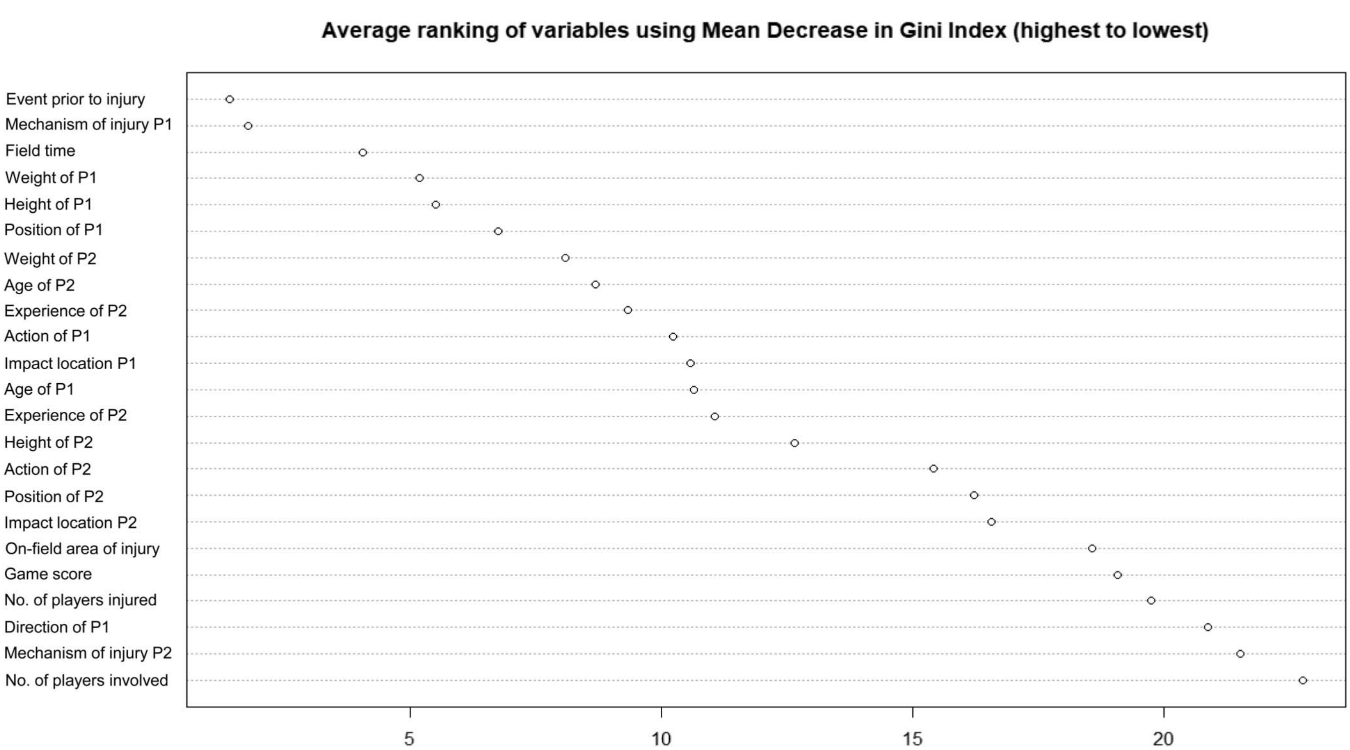

Figure 1. The average importance ranking of the variables over all CV folds (higher rank is equivalent to more importance) in the male dataset. P1 and P2 represents player 1 and player 2, respectively.

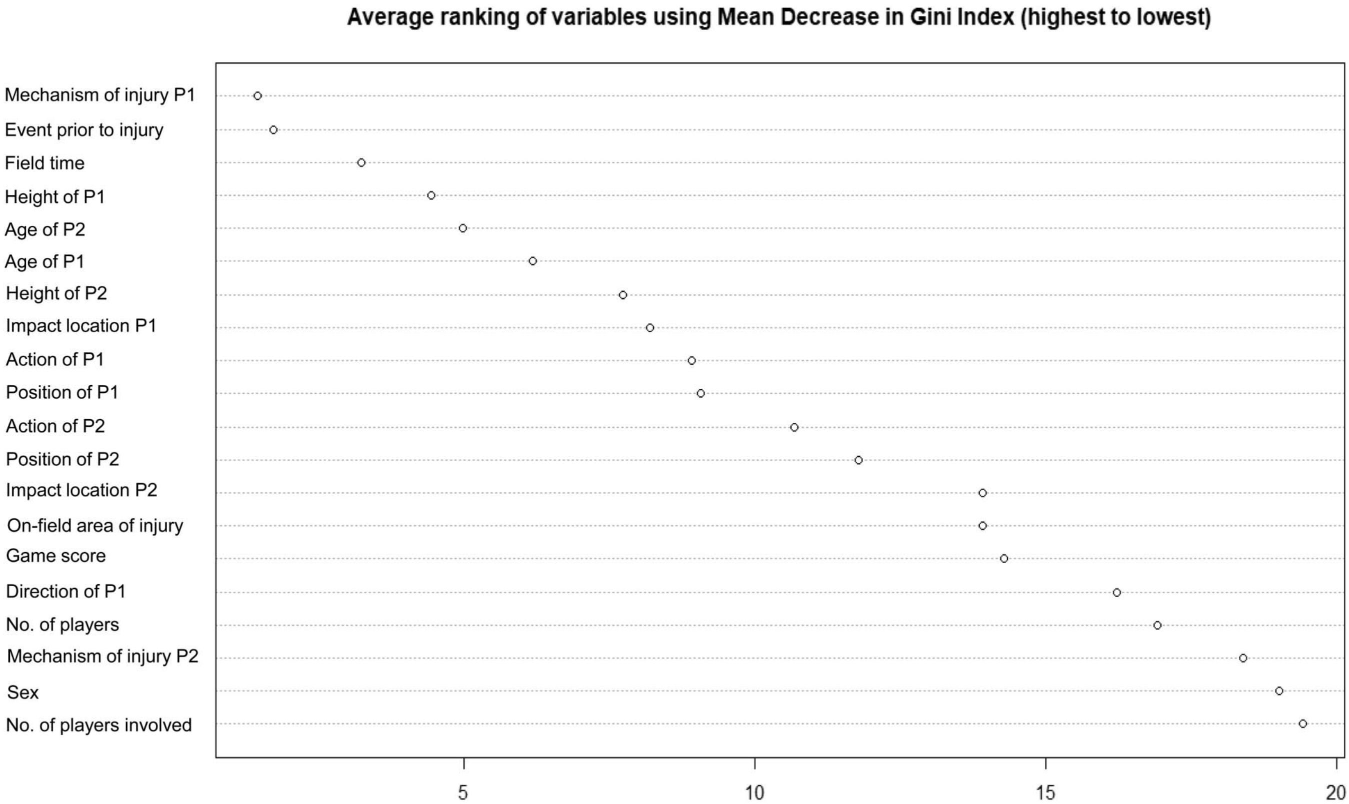

Figure 2. The average importance ranking of the variables over all CV folds (higher numerical rank is equivalent to more importance) in the mixed dataset. P1 and P2 represents player 1 and player 2, respectively.

with the severity of events (Fisher exact test, P-value 0.000424). Either of these 2 preinjury events were present in 11.8% (31/263) HCEs and in 28.9% (13/45) of all severe HCEs. In 42% (13/31) of HCEs that had "Corner kick" or "throw-in," severe injury occurred. In comparison, only 13.8% (32/232) of the remaining HCEs were severe. However, with respect to mechanism of injury, head-to-head and knee-to-head together were not significantly associated (Fisher exact test, P-value = 0.1113) with severity. As noted from the P-value, it is not very far from the level-of-significance (P = 0.05); these 2 mechanisms were present in 31.1% (14/45) severe HCEs and in 21.7% (57/263) of all HCEs. In variable importance for predicting severity, sex was the next-to-last important variable.

DISCUSSION

Main Findings

The ML models were able to classify and distinguish between severe and nonsevere events based on preinjury variables. Based on the AUC, the male dataset performed better than the mixed dataset, a potential consequence of having more variables or a possible indication of differences between characteristics of HCEs between women's and men's soccer.

The potential differences between men's and women's soccer are also highlighted by the fact that in the mixed dataset, the mechanisms "head-to-head" and "knee-to-head" together were no longer significantly associated with severity as they were in the male dataset. However, our data show that the most common mechanisms of HCEs in men's soccer, "head-to-head" and "elbow-to-head," are also the most common mechanisms for HCEs in women's soccer. This indicates that there is little difference between men and women's soccer in mechanisms of HCEs, but potential differences in the mechanisms that lead to severe HCEs. In the future, it would be informative to have different models for male and women's soccer. We did not perform a separate modeling in the women's dataset because only 6 severe HCEs occurred. However, it must be noted that the model still found the variable mechanism of injury as important for predicting severity in the male and mixed dataset.

For both datasets, the event before injury was an important variable for predicting severity. The specific events "corner kicks" and "throw-ins" were together significantly associated with severity of HCEs in both datasets. It is important to note, that unlike the mechanisms of injury that were associated with severity, "corner kicks" and "throw-ins" were not the most common events for HCEs. This indicates that although these events do not occur often before an HCE, when they do occur, they are more likely to cause a more severe HCE. Thus, HCE occurring after corner kicks and throw-ins should particularly warrant attention by officials and formal medical assessment.

Compatibility With Existing Research

Our findings are consistent with previous research in soccer that found that head-to-head contact between players is one of the most common mechanisms of head injury.[23,24] Biomechanical investigations into mechanisms of head injuries in soccer found that head-to-head contact increased risk of concussion compared with other mechanisms.[25,26] Research on HCEs in soccer found that the most common region for head impact location is the frontal region, parietal and occipital region, and temporal region.[21] The high incidence of head-to-head contact in soccer paired with its increased risk of concussion contributes to it being associated with severity.

Research has shown that most head injuries in soccer occur while both players are jumping and attempting to head the ball.[27,28] This is consistent with our findings because corner kicks and throw-ins in soccer result almost exclusively in the ball being put into play aerially, increasing the likelihood of players challenging for the ball with their head as the focal point. Our findings are compatible with the existing literature on the mechanisms of head injuries in soccer while also further investigating their impact on severity.

Applications of Machine Learning and Further Directions

The approach of using ML models to predict the severity of HCEs based on preinjury factors creates an opportunity for improving the detection and treatment of SRC. As opposed to traditional parametric linear modeling, which requires explicitly stating variables, nonlinearities, and interactions; ML techniques such as random forest offer a nonlinear and nonparametric route to learn the interactions from the data based on the response variable.[29] Moreover, random forest can handle high-dimensional data and is found to perform well in presence of class-imbalance as well.[30] In our work, we obtained a good predictive performance and a set of variables important for predicting severity on a small dataset with 20 or more variables based on the dataset we used.

Per the important variables returned by the model, on-field medical personnel should have a heightened awareness regarding mechanisms that involve head-to-head or knee-to-head contact and those occurring from corner kicks or throw-ins. Players injured in these circumstances warrant particularly heightened attention for medical assessment because they are more likely to predict severe events. A better understanding of the factors that result in SRC would allow for the implementation of changes that could be applied by referees and medical personnel to protect athlete health. Our study gives guidance into how modifications could be made to improve the safety of players. For example, referees can be trained that when players sustain head-to-head contact, especially after corner kicks and throw-ins, play should automatically be stopped for medical assessment. In addition, Video Assistant Referees at elite levels of soccer can identify these events and direct the on-field referee to stop play for assessment. Along the same vein, referees could also be directed to be more assertive with fouls committed during corner kicks and throw-ins to protect players when making aerial challenges. By strictly enforcing the rules, players may be deterred from making potentially dangerous challenges during aerial duels. As evident by the recent introduction of additional permanent concussions substitutions in some professional leagues, stakeholders are willing to modify rules to protect player health.[31,32] These findings provide valuable information for soccer authorities to recognize events and mechanisms that increase the severity of HCEs and to aid in the implementation of preventative measures. Our study demonstrates the capabilities of ML in further aiding the understanding of health issues in sports.

Limitations

As is the nature of ML, our predictions and external validation will improve as more data are obtained. In the future studies, one could include factors such as referee history, match importance, and injury history data to increase the accuracy of prediction. In addition, our small sample size for women's data precluded an independent analysis for women's soccer, but future work should investigate potential differences in HCE characteristics further. As more data are collected, future ML research has the ability not just to predict the severity of an HCE, but predict the occurrence of HCEs. Our data are also limited by its dependence on the severity of the HCE based on signs exhibited by the player on video analysis and not on any medical diagnosis. Although our ratings of HCEs had good reliability, camera angles and capture could have been limited for certain factors. Finally, a caution that data from elite-international soccer may not be generalizable to all levels of the game.

CONCLUSIONS

Machine learning was able to use preinjury factors to predict the severity of HCEs in elite soccer and showed that head-to-head and knee-to-head mechanism, along with HCEs occurring from corner kicks and throw-ins, are associated with the severity of HCEs. This provides important information that medical personnel, referees, and league officials can adopt to initiate, and, where needed, institute new and proper strategies to optimize medical assessment and treatment of players.

References

1. Harmon KG, Drezner J, Gammons M, et al; American Medical Society for Sports Medicine. American Medical Society for Sports Medicine position statement: concussion in sport. *Clin J Sport Med*. 2013;23:1–18.
2. National Research Council, Committee on Sports-Related Concussions in Youth. *Sports-Related Concussions in Youth: Improving the Science, Changing the Culture*. Washington, DC: National Academies Press; 2014.
3. Agel J, Evans TA, Dick R, et al. Descriptive epidemiology of collegiate men's soccer injuries: national collegiate athletic association injury surveillance system, 1988–1989 through 2002–2003. *J athletic Train*. 2007;42:270–277.
4. Covassin T, Swanik CB, Sachs ML. Sex differences and the incidence of concussions among collegiate athletes. *J athletic Train*. 2003;38:238.
5. Tarzi C, Aubrey J, Rotundo M, et al. Professional assessment of potential concussions in elite football tournaments. *Inj Prev*. 2020;26:536–539.
6. Cusimano MD, Casey J, Jing R, et al. Assessment of head collision events during the 2014 FIFA World Cup tournament. *Jama*. 2017;317: 2548–2549.
7. Abraham KJ, Casey J, Subotic A, et al. Medical assessment of potential concussion in elite football: video analysis of the 2016 UEFA European championship. *BMJ Open*. 2019;9:e024607.
8. Tarzi C, Tarzi G, Walker M, et al. Medical assessment of head collision events in elite women's and men's soccer. *JAMA*. 2020;323:275–276.
9. Sato M, Morimoto K, Kajihara S, et al. Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma. *Scientific Rep*. 2019;9:7704–7707.
10. Ning Y, Muthiah S, Rangwala H, et al. Modeling precursors for event forecasting via nested multi-instance learning. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. August 13, 2016. pp. 1095–1104.
11. Claudino JG, de Oliveira Capanema D, de Souza TV, et al. 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;5:1–2.
12. Adetiba E, Iweanya VC, Popoola SI, et al. Automated detection of heart defects in athletes based on electrocardiography and artificial neural network. *Cogent Eng*. 2017;4:1411220.

13. Ertelt T, Solomonovs I, Gronwald T. Enhancement of force patterns classification based on Gaussian distributions. *J Biomech*. 2018;67: 144–149.
14. Pensgaard AM, Ivarsson A, Nilstad A, et al. Psychosocial stress factors, including the relationship with the coach, and their influence on acute and overuse injury risk in elite female football players. *BMJ Open Sport Exerc Med*. 2018;4:e000317.
15. López-Valenciano A, Ayala F, Puerta JM, et al. A preventive model for muscle injuries: a novel approach based on learning algorithms. *Medicine Sci Sports Exerc*. 2018;50:915–927.
16. Rossi A, Pappalardo L, Cintia P, et al. Effective injury forecasting in soccer with GPS training data and machine learning. *PloS One*. 2018;13: e0201264.
17. Jaspers A, De Beéck TO, Brink MS, et al. Relationships between the external and internal training load in professional soccer: what can we learn from machine learning? *Int J Sports Physiol Perform*. 2018;13: 625–630.
18. López-Valenciano A, Ruiz-Pérez I, Garcia-Gómez A, et al. Epidemiology of injuries in professional football: a systematic review and meta-analysis. *Br J Sports Med*. 2020;54:711–718.
19. Breiman L. Random forests. 10.1023. A: 1010933404324. 2001.
20. Team RC. R: A language and environment for statistical computing.
21. Chen RC, Dewi C, Huang SW, Caraka RE. Selecting critical features for data classification based on machine learning methods. *J Big Data*. 2020; 7:1–26.
22. Zhu M, Xia J, Jin X, et al. Class weights random forest algorithm for processing class imbalanced medical data. *IEEE Access*. 2018;6: 4641–4652.
23. Armstrong N, Rotundo M, Aubrey J, et al. Characteristics of potential concussive events in three elite football tournaments. *Inj Prev J Int Soc Child Adolesc Inj Prev*. 2020;26:334–338.
24. Andersen TE, Arnason A, Engebretsen L, Bahr R. Mechanisms of head injuries in elite football. *Br J Sports Med*. 2004;38:690–696.
25. Withnall C, Shewchenko N, Wonnacott M, Dvorak J. Effectiveness of headgear in football. *Br J Sports Med*. 2005;39(suppl 1):i40–i48; discussion i48.
26. Withnall C, Shewchenko N, Gittens R, Dvorak J. Biomechanical investigation of head impacts in football. *Br J Sports Med*. 2005; 39(suppl 1):i49–i57.
27. Fuller CW, Junge A, Dvorak J. A six year prospective study of the incidence and causes of head and neck injuries in international football. *Br J Sports Med*. 2005;39(suppl 1):i3–i9.
28. Beaudouin F, Aus der Fünten K, Tröß T, et al. Match situations leading to head injuries in professional male football (Soccer)—a video-based analysis over 12 years. *Clin J Sport Med*. 2020;30:S47–S52.
29. Grömping U. Variable importance assessment in regression: linear regression versus random forest. *Am Statistician*. 2009;63:308–319.
30. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. *BMC Med Inform Decis Making*. 2011;11:51–53.
31. *Concussion substitutions: What you need to know [Internet]. Premier League Football News, Fixtures, Scores & Results*. 2021. Available at: https://www.premierleague.com/news/1980769. Accessed December 15, 2021.
32. Tarzi G, Tarzi C, Mirsu D, et al. Effect of a new concussion substitute rule on medical assessment of head collision events in Premier League football. *Inj Prev*. 2022. [epub ahead of print].