International Journal of Sports Science and Physical Education
2024, Vol. 9, No. 3, pp. 40-46
https://doi.org/10.11648/j.ijsspe.20240903.11

Research Article

Forecasting of Injuries in Ethiopia Premier League: Time Series Model Analysis

Aychew Kassa Belete[1, *], Fasiledes Fetene Asfaw[2], Birhan Ambachew Taye[2], Bantie Getinet Yirsaw[3]

[1]Department of Sport Science, Faculty of Natural and Computational Science, Woldia University, Woldia, Ethiopia
[2]Department of Statistics, Faculty of Natural and Computational Science, Woldia University, Woldia, Ethiopia
[3]Department of Epidemiology of Biostatistics, University of Gondar, Gondar, Ethiopia

Abstract

Background: Sport injury is an injury which is occurs in playing field maybe in training or competition. Epidemiology of sports injury on male footballer has been documented that injury incidences were 10-35 injuries per 1000 game hours. The main objective of our study is predicting the number of injuries for coming specific time by analyzing historical injury data obtained from team physicians. *Methods*: We collected historical injury data from the Ethiopia Premier League which is collected for 50 weeks, including the number of injuries, types of injuries, affected players, and duration of absence from play. We then selected an appropriate time series model for forecasting injuries based on the nature of the data and its patterns, considering potential models such as ARIMA (AutoRegressive Integrated Moving Average). After training the selected time series model using historical injury data and validating its performance by comparing predicted values with actual injury occurrences, we used it to forecast injuries for the upcoming seasons of the Ethiopia Premier League. *Results*: In Ethiopia the weekly average increment in sport injury from week 1to week 50 was 4.4. The maximum number of sport injury occurred on week 30. The series is not stationary at level, but the series is stationary at first difference. The selected model in this study was ARIMA (3,0,0) that has small AIC and BIC. Based on ARIMA (3,0,0) model the new sport injury in Ethiopia premier league was 13 injuries in week 51, and the forecasted number of injuries for the following weeks were 12, 12, 13, 11, 11, 10, 10, 10, and 10, respectively, up to week 60. *Conclusion*: our research finding indicates that, occurrence of sport injury will increase for coming weeks so that teams should implement injury prevention programs, prioritize rest and recovery, and ensure access to qualified medical staff for immediate care and rehabilitation.

Keywords

Sport Injury, Forecasting, Football, Autoregressive Integrative Moving Average

1. Introduction

Sport injuries are common occurrences that can happen to athletes of all levels and ages [1]. These injuries can range from minor sprains and strains to more serious fractures and tears [2]. They can occur during any type of physical activity,

*Corresponding author: ayichew.k@wldu.edu.et (Aychew Kassa Belete)

Received: 14 May 2024; **Accepted:** 12 June 2024; **Published:** 6 August 2024

such as running, jumping, or contact sports. Sport injuries can have a significant impact on an athlete's performance and overall well-being, requiring proper treatment and rehabilitation to ensure a full recovery [3-5]. Understanding the frequency, causes, prevention, and treatment of sport injuries is essential for athletes, coaches, and healthcare professionals alike [6].

In the world of football, sport injuries are a common occurrence due to the physical nature of the game [7]. Some of the most common injuries in football include sprains and strains, knee injuries, concussions, and fractures. These injuries can occur during tackles, collisions, or simply from overexertion during training and games [8]. The causes of these injuries can vary, but often stem from factors such as poor technique, inadequate warm-up and stretching, fatigue, or contact with other players [9, 10]. Additionally, playing on hard surfaces or in adverse weather conditions can also increase the risk of injury. Prediction is key when it comes to minimizing the risk of sport injuries in football [11, 12].

It is difficult to forecast the exact number of sport injuries in football, as it can vary based on a number of factors such as player conditioning, game intensity, and playing conditions [13]. However, it is likely that sprains, strains, knee injuries, concussions, and fractures will continue to be common injuries in football due to the physical nature of the sport. But, some scholars suggest that we can forecast sport injuries by taking into account previous history [14]. By analyzing the types and frequency of injuries that have occurred in a particular sport, researchers can identify patterns and trends that may indicate areas of concern. This information can then be used to develop targeted injury prevention strategies and protocols to help minimize the risk of future injuries. By understanding the common types of injuries that occur in a sport, coaches, players, and medical staff can be better prepared to address and mitigate potential risks, ultimately promoting a safer and healthier environment for athletes.

Epidemiology of sports injury on male footballer has been documented that injury incidences were 10-35 injuries per 1000 game hours. In overall rate of injuries in contact sports were assessed that 10 to 15 injuries per 1000 playing hours [15]. Due to their effects on individuals' emotional states and a team's performance, professional athletes' injuries have a significant effect on the sports industry. Additionally, the expense of a player's recovery and rehabilitation is frequently high, both in terms of medical care and lost income resulting from the player's fame [16]. Recent research demonstrates that injuries in Spain cause about 16% of season absence by professional soccer players, corresponding to a cost of around 188 million euros per season [17].

There is a lack of published research on predicting sports-related injuries in the Betking Ethiopian Premier League. While some studies have addressed the prevalence and incidence of such injuries, there is a historical pattern of football injuries in the national team, clubs, and other league players in Ethiopian football. These injuries have caused significant problems for players, particularly those in the premier league. Recent research by Adise Alemu has shown an increase in the number of injuries in the Ethiopia Premier League. Despite the growing concern for athlete safety and injury prevention in sports, there is a lack of comprehensive data and analysis on the types and frequency of injuries occurring in the Ethiopia Premier League. This hinders the ability to effectively forecast and address potential injury risks, leading to a higher likelihood of preventable injuries among athletes. Therefore, there is a need to develop a time series model analysis to forecast sport injuries in the Ethiopia Premier League, in order to improve injury prevention strategies and promote a safer and healthier environment for athletes.

2. Data and Methods

Source of Data

Historical injury data collected from the Betking Ethiopian Premier League by distributing questionnaires to all sports medicine specialists working with the clubs at the beginning of the tournament. Prior to the competition, convened all the sports doctors and had a discussion, urging them to document any type of injury, ranging from mild to severe, including abrasions, swelling, dislocations, and fractures, as well as the number of injuries, types of injuries, affected players, and duration of absence from play.

3. Methods

Historical injury data were collected from the Ethiopia Premier League, including the number of injuries, types of injuries, affected players, and duration of absence from play. Then selected an appropriate time series model for forecasting injuries based on the nature of the data and its patterns, considering potential models such as ARIMA (AutoRegressive Integrated Moving Average). After training the selected time series model using historical injury data and validating its performance by comparing predicted values with actual injury occurrences, and forecast injuries for the upcoming seasons of the Ethiopia Premier League. This provided estimates of the expected number of injuries, their distribution, and potential impact on teams and players. Finally, analyzed the forecasted injury data and provided recommendations for injury prevention strategies, player management, and medical support within the league. Then added up the amount of injuries found in each group then, divided them into 50 weeks.

3.1. Stationary Assumption

A stationary process has the property that the first and second moments do not change over time. It is an essential property to define a time series process. Time series, X_t is a finite variance process such that (i) the mean value function, μ_t is constant and does not depend on time t, and the auto covariance function, $\gamma(s, t)$ depends on s and t only through their deference $|s\text{-}t|$.

3.2. Model Specification

The word ARIMA stands for Auto-Regressive Integrated Moving Average. ARIMA models are the most general class of models for forecasting a time series data y_t.[3] The ARIMA forecasting equation for a stationary time series is a linear equation in which the predictors consist of lags of the dependent variable and/or lags of the forecast errors. The general ARIMA (p, d, q) model can be written:

4. Results

There are two ways of testing stationarity these are graphical and common test. In the research graphical methods included time series plot and common tests include Augmented Dickey-Fuller and Philips-Perron test statistic.

Table 1. *Descriptive statistics.*

	Quantities of injuries
Mean	4.40
Median	4.50
Std. Dev	1.54474
Minimum	1
Maximum	9
Sum	132

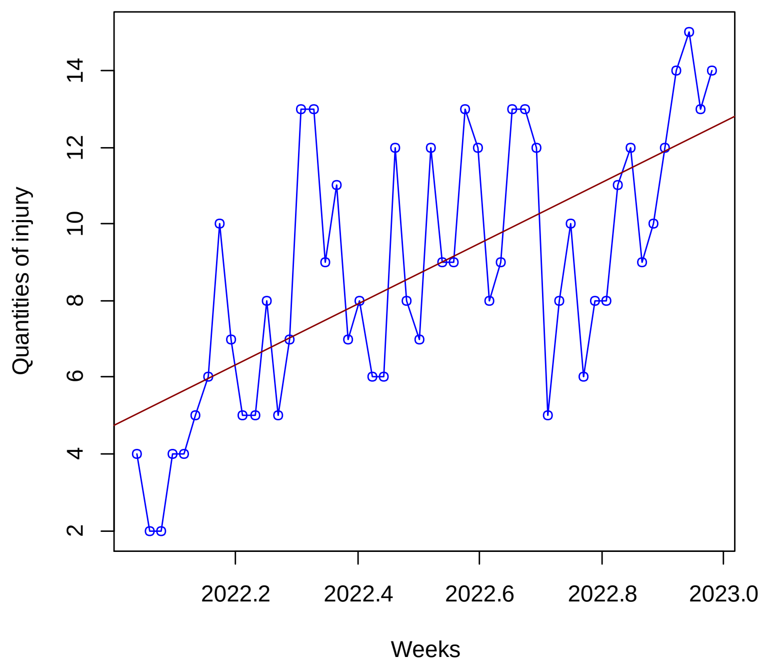

Figure 1. *Time series plot of Quantity of Injury.*

From figure above the mean and variances of the variable seems increase with time, implies the series is nonstationary. Since the decision is subjective, the stationary process also cannot be determined from above figure. Therefore, it is better to undertake a formal test for unit root test.

Grounded on the table below the Augmented Dickey-Fuller test p-value was greater than the level of significance alpha ($0.1149>0.05$) implies at 5% level of significance we have evidence not to reject the null hypothesis therefore before differencing the data the researcher used log to base ten transformation of the data then tested stationarity. The transformed data is stationary ($0.04345<0.05$).

Table 2. *Test of Stationary for Injury.*

Tests	**ADF at level**			
	Quantities of injury at level		**Log(quantities)**	
Quantities	Test statistics	p-value	Test statistics	p-value
	-3.5831	0.1449	-3.0704	0.04345

5. Model Selection

In univariate time series analysis the most known and applicable family of model is ARIMA(p,d,q). ARMA stands for autoregressive integrated moving average. This model is joint model of two parts autoregressive (AR) and moving average (MA). The General ARIMA(p,d,q) model is written as

$$\gamma(B)(1-B)^d Y_t = \theta(B)e_t$$

Where B is the back shift operator $BY_t = Y_{t-1}, \gamma(B) = (1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$ and $\theta(B) = 1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_q B^q$.

Before proceeding to select the model, it is better to look out the autocorrelation and partial autocorrelation plot. These two plots give as the clue on how many lags should be incorporated on the AR and MA part of the ARIMA model. The ACF helps to identify the MA part and the PACF identifies the AR part. The figures below show the autocorrelation functions and partial autocorrelation for the differenced series. Values above the line are significant. (The height of the dashed line is determined by the amount of data.) From the figure below the differenced series of the two plots, the ACF plot cuts at lag three and the PACF plot at lag one therefor, the first candidate model is ARMA (3,0,1).

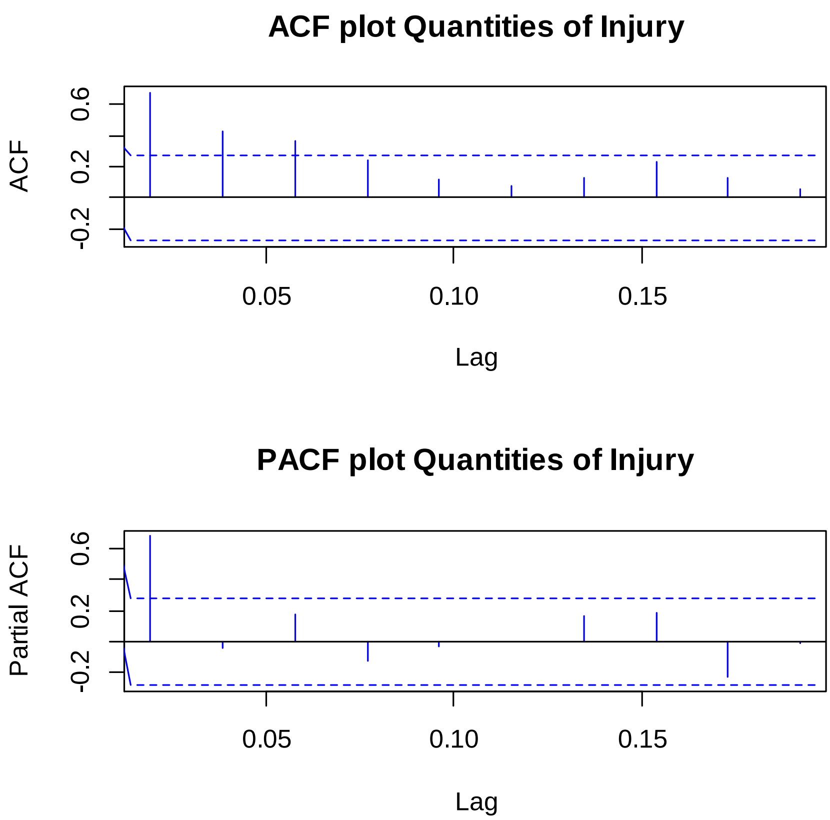

Figure 2. *ACF and PACF Plot for Injury.*

There are many candidates model for the differenced series. Among these candidates the best model was selected using Akaike information criteria (AIC). From the table below the best model selected using R version 4.2.1 with minimum AIC, was ARIMA (3,0,0) with non-zero mean with AIC value 29.92. which can be written as

$$\ln(y_t) = \alpha + \phi_1 lny_{t-1} + \phi_2 lny_{t-2} + \phi_3 lny_{t-3} + \varepsilon_t$$

Table 3. Model Selection criteria.

Model	AIC*
ARIMA(2,0,2)	INF
ARIMA(0,0,0)	72.33
ARIMA(1,0,0)	40.95
ARIMA(0,0,1)	50.55
ARIMA(2,0,0)	35.89
ARIMA(3,0,0)	29.92*
ARIMA(4,0,0)	34.61
ARIMA(3,0,1)	33.82
ARIMA(2,0,1)	INF
ARIMA(4,0,1)	38.41

5.1. Parameter Estimation

Table below shows the significance of the autoregressive components. AR(1) with p-value 0.0001, AR(3) with p-value 0.02197 and intercept term with p-value 0.0001 are significant but the AR(2) with p-value 0.4995 coefficient is constrained to be zero.

Table 4. Parameter Estimated.

Coefficients:	Estimate	Std.error	z-value	Pr(>|Z|)	2.5 %	97.5 %
ar1	0.6484	0.1349	4.8064	0.0001***	0.384	0.913
ar2	-0.11194	0.1658	-0.6752	0.4995	-0.437	0.213
ar3	0.3409	0.1488	2.2909	0.02197*	0.049	0.633
Intercept	0.87572	0.13074	6.6984	0.0001***	1.426	2.606

Signcodes: '***' 0.001 '**' 0.01 '*' 0.05

5.2. Model Diagnostic

At this stage, an evaluation of the tentative model based on the estimated residual properties is to be performed.

5.3. Normality of the Residual

In this research we use quantile-quantile plots and Shapiro-Wilk normality test on standardized residuals to assess normality. First of all it could be useful to examine the standardized residual plot. Quantile-Quantile plots are used to assess whether the data in a single series are normally distributed or not. If the two distributions are the same, the QQ-plot should lie on a straight line. Based on the graph and table below, the QQ plot and the Shapiro-Wilk normality test (W = 0:96406, p-value= 0.1315) would not lead us to reject normality of the error terms in this model.

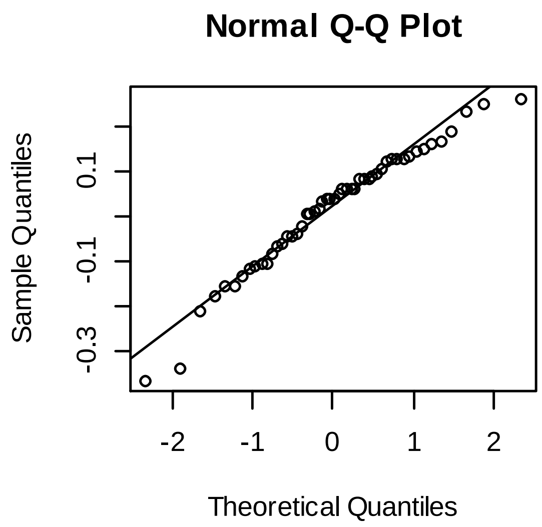

Figure 3. Q-Q Plot of the residuals.

Table 5. *Test of Normality.*

Shapiro-Wilk normality test	W	P-value
data: residuals (model)	0.96406	0.1315

5.4. Autocorrelation of Residuals

From the graph and table below p-values for Ljung-Box is 0.2044. Since p-value is large we can conclude that there is no autocorrelation between residuals. Therefore, there is no evidence of autocorrelation in the residuals of this model. The selected model is an appropriate model. Forecasting using ARIMA (3,0,0).

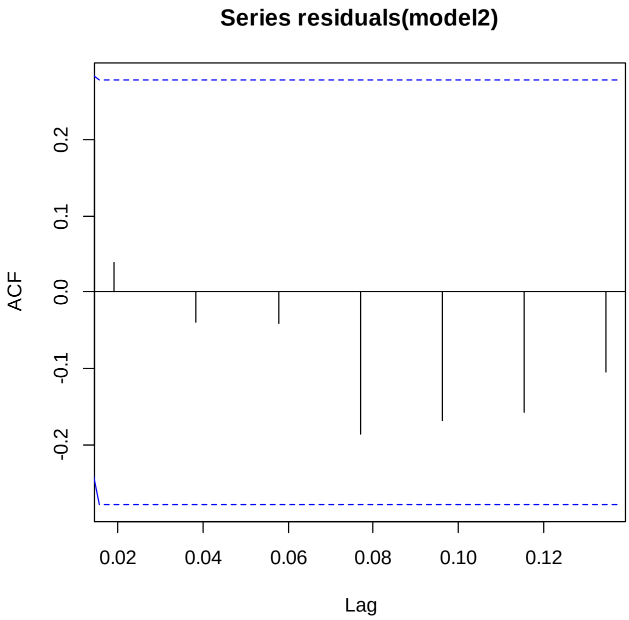

Figure 4. *ACF Plot of the residuals.*

Table 6. *ACF Plot of the residuals.*

Box-Ljung test	X-squared	df	p-value
Model	5.9301	4	0.2044

Therefore, in this research the selected and best model for the data is given by:

$$\ln(y_t) = 0.87572 + 0.64837 lny_{t-1} + 0.34092 lny_{t-3} + \varepsilon_t$$

5.5. Forecasting the New Quantities of Injuries

One of the major applications of time series analysis is forecasting the future values based on the model build up.

Table 7. *Forecasted values.*

WEEK	Point forecast	Lo 95	Hi 95
51	13.38995	0.8503475	1.403211
52	12.28694	0.7599924	1.418895
53	11.97773	0.7380690	0.7380690
54	11.71604	0.7046828	1.432879
55	11.2479	0.6625788	1.439564
56	10.88653	0.6351210	1.438658
57	10.62.689	0.6134717	1.439340
58	10.35521	0.5910836	1.439234
59	10.0973	0.5713272	1.437083
60	9.880716	0.5549356	1.434641

6. Conclusion

Based on previous history, it is likely that football will continue to see a high number of sprains, strains, knee injuries, concussions, and fractures. However, with advancements in sports science and injury prevention techniques, there is potential for a reduction in the overall number and severity of injuries. It will be important for players, coaches, and medical staff to remain proactive in addressing and treating injuries as they occur, while also promoting a culture of safety and injury prevention within the sport.

The ARIMA (3,0,0) model was chosen for this study due to its low AIC and BIC values. According to the model, there were 13 injuries in week 51, and the forecasted number of injuries for the following weeks were 12, 12, 13, 11, 11, 10, 10, 10, and 10, respectively, up to week 60. This indicates a rising trend in sports injuries over the next 10 weeks. Therefore, for the future athlete's dream to come true, the relevant body is expected to do a lot of work by understanding the frequency and causes of injuries found by other researchers. There are multiple factors contributing to this trend, such as lack of awareness, unlicensed sports doctors, inadequate knowledge about active rest, insufficient use of proper equipment, and a playing philosophy that still heavily relies on force.

This research result prevails that, the most frequent injury in the Ethiopian Premier League is muscle strains, particularly in the lower extremities. This is often attributed to the fast-paced nature of the game and the high physical demands placed on the players. Additionally, ankle and knee injuries are also common due to the frequent changes in direction and sudden stops and starts during matches. Concussions and head injuries are also a concern, especially during aerial challenges and collisions. Overall, the physical nature of the sport and the intensity of competition contribute to a high incidence of various types of injuries in the Ethiopian Premier League.

To prevent injuries in football, it is crucial for players to engage in proper warm-up and cool down routines, focus on strength and conditioning training, and emphasize proper technique during training sessions. Additionally, teams should implement injury prevention programs, prioritize rest and recovery, and ensure access to qualified medical staff for immediate care and rehabilitation. Player education on the importance of injury prevention and reporting any discomfort or pain is also essential to prevent minor issues from developing into more serious injuries. By implementing these strategies, teams can help reduce the risk of injuries and keep players healthy and performing at their best.

Abbreviations

ARIMA	Auto Regressive Integrated Moving Average
ACF	Auto Correlation
AIC	Akaike Information Criteria

Acknowledgments

Thank you very much the data collection teams and study participants who made this study possible.

Author Contributions

Aychew Kassa Belete: Concept, analysis, design and interpretation

Fasiledes Fetene Asfaw: Develops the model

Birhan Ambachew Taye: Data collection and analysis

Bantie Getnet Yirsaw: Data collection and analysis

Data Availability Statement

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interests.

References

[1] Kujala, U. M., et al., Acute injuries in soccer, ice hockey, volleyball, basketball, judo, and karate: analysis of national registry data. Bmj, 1995. 311(7018): p. 1465-1468.

[2] Maffulli, N., et al., Sport injuries: a review of outcomes. British medical bulletin, 2011. 97(1): p. 47-80.

[3] Parsons, J. T. and A. R. Snyder, Health-related quality of life as a primary clinical outcome in sport rehabilitation. Journal of sport rehabilitation, 2011. 20(1): p. 17-36.

[4] Taylor, J., et al., Comprehensive sports injury management: From examination of injury to return to sport. 2003: Lippincott Williams & Wilkins.

[5] Smith, A. M., S. G. Scott, and D. M. Wiese, The psychological effects of sports injuries coping. Sports Medicine, 1990. 9: p. 352-369.

[6] Saragiotto, B. T., C. Di Pierro, and A. D. Lopes, Risk factors and injury prevention in elite athletes: a descriptive study of the opinions of physical therapists, doctors and trainers. Brazilian journal of physical therapy, 2014. 18: p. 137-143.

[7] Gurau, T. V., et al., Incidence and prevalence of injuries in some sports-review. Balneo & PRM Research Journal, 2023. 14(4).

[8] Olson, D., et al., Injuries in professional football: current concepts. Current sports medicine reports, 2013. 12(6): p. 381-390.

[9] Bahr, R. and S. Mæhlum, Clinical guide to sports injuries. 2004: Human Kinetics.

[10] Olson, D. E., et al., Football injuries: current concepts. Current sports medicine reports, 2011. 10(5): p. 290-298.

[11] Chivers, I. Turfgrass sports surfaces and their relationship to player injuries. in II International Conference on Turfgrass Science and Management for Sports Fields 783. 2007.

[12] Gökçe, U., et al., POTENTIAL RISKS OF OUTDOOR SPORTS IN A CHANGING CLIMATE. SPORMETRE Beden Eğitimi ve Spor Bilimleri Dergisi, 2022. 20(1): p. 145-163.

[13] Inklaar, H., Soccer Injuries: II: Aetiology and Prevention. Sports medicine, 1994. 18: p. 81-93.

[14] Gilbert, F. C., et al., Association between concussion and lower extremity injuries in collegiate athletes. Sports health, 2016. 8(6): p. 561-567.

[15] López-Valenciano, A., et al., Epidemiology of injuries in professional football: a systematic review and meta-analysis. British journal of sports medicine, 2020. 54(12): p. 711-718.

[16] Baillie, P. H. and S. J. Danish, Understanding the career transition of athletes. The sport psychologist, 1992. 6(1): p. 77-98.

[17] Rossi, A., et al., Effective injury prediction in professional soccer with GPS data and machine learning. stat, 2017. 1050: p. 23.