Original Research

Development of an Injury Burden Prediction Model in Professional Baseball Pitchers

Garrett Bullock[1,2] [a], Charles Thigpen[3,4], Gary Collins[5,6], Nigel Arden[1,2], Thomas Noonan[7,8], Michael Kissenberth[9], Ellen Shanley[3,4]

[1] Centre for Sport, Exercise and Osteoarthritis Research Versus Arthritis, University of Oxford, [2] Department of Orthopaedic Surgery & Rehabilitation, Wake Forest University School of Medicine, [3] University of South Carolina Center for Rehabilitation and Reconstruction Sciences, [4] ATI Physical Therapy, [5] Centre for Statistics in Medicine, University of Oxford, [6] Oxford University Hospitals NHS Foundation Trust, [7] Department of Orthopaedic Surgery, University of Colorado School of Medicine, [8] University of Colorado Health, Steadman Hawkins Clinic, [9] Steadman Hawkins Clinic of the Carolinas

International Journal of Sports Physical Therapy

Background

Baseball injuries are a significant problem and have increased in incidence over the last decade. Reporting injury incidence only gives context to rate but not in relation to severity or injury time loss.

Hypothesis/Purpose

The purpose of this study was to 1) incorporate both modifiable and non-modifiable factors to develop an arm injury burden prediction model in Minor League Baseball (MiLB) pitchers; and 2) understand how the model performs separately on elbow and shoulder injury burden.

Study Design

Prospective longitudinal study

Methods

The study was conducted from 2013 to 2019 on MiLB pitchers. Pitchers were evaluated in spring training arm for shoulder range of motion and injuries were followed throughout the season. A model to predict arm injury burden was produced using zero inflated negative binomial regression. Internal validation was performed using ten-fold cross validation. Subgroup analyses were performed for elbow and shoulder separately. Model performance was assessed with root mean square error (RMSE), model fit (R2), and calibration with 95% confidence intervals (95% CI).

Results

Two-hundred, ninety-seven pitchers (94 injuries) were included with an injury incidence of 1.15 arm injuries per 1000 athletic exposures. Median days lost to an arm injury was 58 (11, 106). The final model demonstrated good prediction ability (RMSE: 11.9 days, R^2: 0.80) and a calibration slope of 0.98 (95% CI: 0.92, 1.04). A separate elbow model demonstrated weaker predictive performance (RMSE: 21.3; R^2: 0.42; calibration: 1.25 [1.16, 1.34]), as did a separate shoulder model (RMSE: 17.9; R^2: 0.57; calibration: 1.01 [0.92, 1.10]).

a **Corresponding Author:**
Garrett S. Bullock PT, DPT, DPhil
Wake Forest School of Medicine
475 Vine St
Winston-Salem, NC 27701
gbullock@wakehealth.edu

Conclusions

The injury burden prediction model demonstrated excellent performance. Caution should be advised with predictions between one to 14 days lost to arm injury. Separate elbow and shoulder prediction models demonstrated decreased performance. The inclusion of both modifiable and non-modifiable factors into a comprehensive injury burden model provides the most accurate prediction of days lost in professional pitchers.

Level of Evidence

2

INTRODUCTION

Baseball injuries are a significant problem with increased incidence over the last decade.[1–3] The greatest injury incidence is to the elbow and shoulder,[1,4] with orthopaedic surgery signifying a significant proportion of injuries.[5] These injuries have a substantial medical,[6] financial,[7] and psychological burden[5] on the individual and team.[8] Due to the significance of these injuries,[6,7] individual physical risk factors have been recognized to assist clinicians in identifying baseball players at risk for arm injury.[9–12] However, these methods are limited by their simplicity, without accounting for the complex interrelated nature of modifiable and non-modifiable relationship to baseball injuries.[13,14]

While identifying individuals at risk for injury can improve athlete injury outcomes,[9] this does not give a complete clinical perspective. Reporting injury incidence only gives the rate but not injury severity or associated time loss.[8] For example, the professional baseball shoulder and elbow injury incidence is 1.4 and 0.9 injuries per 1000 athlete exposures, respectively,[15] but elbow injuries have over six times more days missed compared to shoulder injuries.[16] Simply identifying baseball players at risk for injury would give equal weight to shoulder and elbow injuries and additionally would not distinguish between minor and severe injuries.[8] Severity and site of injury will relate to the requirement for surgical interventions, including reconstruction, and hence time loss, highlighting the need to assess these parameters.[17]

Further, considerations must be given to the unbalanced nature of injury data, as the majority of athletes will report no injuries (i.e., no days lost), with a sparse number sustaining severe injuries, providing greater complexity to injury burden data.[18] Pinpointing potential high injury burden athletes (pitchers) would allow for improved resource consolidation.[8] Due to the continued rise in pitching injuries,[1,2] and the incomplete clinical context of the identified injury risk factors,[16] there is a need to distinguish pitchers at increased risk for greater comprehensive arm injury burden.

Pitching is a series of coordinated movements that involve the entire body.[19] The interplay between modifiable and non-modifiable intrinsic and extrinsic factors in determining injury risk and burden signify the intricacies of this problem.[13] The complex nature of pitching injuries require in depth examination, reasoning, and clinical decisions.[20] However, while sports medicine clinicians have advanced skills and reasoning,[21] there continues to be ambiguity in determining true risk.[22] A more complex approach to identify at risk athletes has been proposed,[23] involving the assessment of modifiable and non-modifiable factors as a whole.[23–25] One method that incorporates this approach is via risk prediction models.[26] Prediction models are statistical models that combine multiple predictors to estimate an individual's risk of an event.[27–29] These tools are not designed to supersede clinical experience and decisions; rather, assist in determining the best actions (or inaction) for patients.[26] Prediction models have been used extensively in clinical medicine, such as whether to prescribe of cholesterol lowering therapies through the Framingham cardiovascular risk score.[30]

Quantifying complex systems through prediction models is needed to enumerate injury burden.[24] These predictions can be used by clinicians and other healthcare professionals to better identify baseball players at risk for a high injury burden,[31] and subsequently improve injury identification.[32] Therefore, the purpose of this study was to 1) incorporate both modifiable and non-modifiable factors to develop an arm injury burden prediction model in Minor League Baseball (MiLB) pitchers; and 2) understand how the model performs separately on elbow and shoulder injury burden.

MATERIALS AND METHODS

STUDY DESIGN

A prospective longitudinal study was conducted from 2013 to 2019 on MiLB pitchers in one Major League Baseball (MLB) organization. Prior to testing, the risk and benefits of participation were explained in verbal and written form to all participants. Participants were tested at the beginning of spring training (pre-season). All test administrators were blinded to hand dominance.[33] Participants were tested for shoulder range of motion (ROM) and humeral torsion (HT). Following testing, participants were followed for the entire season for total athletic exposures and arm injuries. Each pitcher was incorporated into each individual season. If a pitcher played multiple seasons, each season was an individual observation. If a pitcher sustained an injury, the pitcher was no longer included in observation for the next season. All participant information was de-identified and coded into an encrypted centralized database. This investigation received favorable ethics approval from the University Institutional Review Board.

PATIENT PUBLIC INVOLVEMENT

Sports medicine clinicians, performance professionals, and coaches within the organization were included in determining the variables to be collected. Presentations were performed for organizational stakeholders throughout data collection and prediction model development.

PARTICIPANTS

Participants were included if they were able to participate in all practices and competitions and were under a MiLB contract. Participants were excluded if they were currently injured or not participating in all spring training activities, participating in MLB spring training at time of data collection, or signed a professional contract in the middle of the season (e.g., draft, free agent, or international signing).[9]

SHOULDER RANGE OF MOTION AND HUMERAL TORSION

Shoulder external (ER), internal (IR), horizontal adduction (HA) range of motion (ROM) was measured using previously described methods.[34–37] Internal reliability testing demonstrated excellent reliability for shoulder ER (intraclass correlation coefficient (ICC) (2,1) =0.99; Standard error of measure (SEM) = 2.0; ICC (2,K) = 0.99; SEM= 0.95), IR (ICC (2,1)= 0.97; SEM = 2.5; ICC (2,K) = 0.97 SEM= 2.6), and HA (ICC (2,1) = 0.99; SEM =2.59; ICC (2,K) =0.97; SEM = 1.2). Shoulder ER and IR were summed to measure shoulder total ROM (TROM). Humeral torsion (HT) were measured with a 5-MHz transducer (Sonosite Inc., Bothell, WA, USA) with previously described methods.[10] HT testing demonstrated excellent reliability (ICC (2,1) = 0.97; SEM = 2.5; ICC (2,K) = 0.97 SEM= 2.6). Two data collection trials were performed per shoulder and the data collection trials were averaged for analyses.[38] For further explanation of the methods, please refer to the Appendix 1.

EXPOSURE

Baseline (i.e., athlete) exposure was defined as number of days participating in training, practice, or games in a season.[39] Pitching exposure was defined as the total number of pitches within a season.[39]

OUTCOME DEFINITION

An injury was defined as an injury to a tendon, ligament, nerve, muscle, or bone that occurred during any baseball team sponsored activity or event and was followed by at least one day of missed practice or baseball games.[33] If a player was unavailable to play for injury prevention reasons (i.e., has reached league or individually determined pitch or innings count limits), then their absence was not considered as an injury. Injuries were defined by the Orchard Sports Injury Classification system and arms injuries stratified by shoulder/clavicle, upper arm, elbow, and forearm.[40] All other injuries and illness were also recorded, and time loss was not taken into account for overall exposure, nor included in the injury burden analyses. Arm injury burden was defined as the product of incidence and severity.[8] Arm injury burden was quantified as the total number of days lost to arm injury in one baseball season.[8]

Residual inspection demonstrated model instability above 90 days lost to arm injury. Following research team and MLB organizational discussion, time loss to arm injury was truncated to 90 days. In other words, if a player sustained an injury of 120 days, within the model, this would be included at 90 days.

STATISTICAL ANALYSES

All data were investigated for missingness prior to analyses using the R package *naniar*. Missing data was low (3% of entire cohort had one or more missing values), thus complete case analyses were performed. Participant characteristics were described using mean (standard deviation) for continuous normally distributed variables, median (interquartile range) for non-normally distributed continuous variables, and frequencies and percentages for categorical variables. Arm injury incidence was calculated by sum of arm injuries divided by sum of baseball (i.e., athlete) exposures (AE) multiplied by 1000. For expanded statistical analyses description, please refer to Appendix 2.

Linearity was not assumed, and prior to model development, continuous variables were assessed for non-linearity in relation to the outcome of days lost to arm injury. All predictors were observed to have linear relationships.

SAMPLE SIZE CALCULATION

An a priori sample size calculation was performed prior to model development.[41] It should be noted that this sample size method does not explicitly cover zero inflated models. During the study period, a total of 297 pitchers met inclusion criteria, therefore a maximum of 26 parameters (i.e., the number of degrees of freedom) could be included in the development of the injury burden prediction model. The R package *pmsampsize* was used to calculate the required sample size.

MODEL DEVELOPMENT: PRIMARY ANALYSIS

To predict the number of days lost (up to 90 days) to arm injury in professional pitchers a zero inflated negative binomial regression model was utilized.[42] An alternative model, a hurdle model was developed, to demonstrate similar prediction performance (Appendix 3). Predictor variables included were chosen based on the baseball injury prevention literature and included: 1) age, 2) BMI, 3) pitching role (starter versus reliever), 4) seasonal number of pitches, 5) number of pitching appearances in a season, 6) HT difference between dominant and nondominant shoulder (in degrees), 7) dominant shoulder TROM (in degrees), 8) dominant HA (in degrees), 9) lower extremity or trunk injury in the same year, 10) any previous arm injury history, 11) years played professionally, 12) received individualized injury prevention programs, 13) continent of origin, and 14) days practiced and competed in the season (exposure).[9,43–52] Variable selection was performed using elastic

net penalization, using ten-fold cross-validation.[53] Internal validation of the model was performed using ten-fold cross validation.[53] Prediction model performance was assessed by calculating the root mean square error (RMSE), explained variation (R^2), and calibration. Calibration is the agreement between predicted and actual risk.[32] Calibration was assessed by calculating the calibration slope with 95% confidence intervals and graphically plotting the observed values against the predicted values. Subgroup analyses were performed for elbow and shoulder separately. A sensitivity analysis was performed with the inclusion of only modifiable predictors to assist clinicians in understanding risk and treatment options including: (1) seasonal number of pitches, 2) number of pitching appearances in a season, 3) dominant shoulder TROM (in degrees), 4) dominant HA (in degrees), 5) received individualized injury prevention programs, 6) days practiced and competed in the season (exposure). All analyses were performed in R version 3.5.1 (R Core Team (2013). The R package *pscl* was used for zero inflation modelling and *mpath* for elastic net. For full code, please see Appendix 4.

Reporting of this study followed the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) recommendations.[54]

RESULTS

PARTICIPANT CHARACTERISTICS

A total of 297 pitchers were included (age: 23.0 (2.2) years, BMI: 24.8 (2.2), left handed: 21%). A total of 84 pitchers reported an arm injury during data collection. Overall arm injury incidence was 1.15 arm injuries per 1000 AE's, 0.5 elbow injuries per 1,000 AE's, and 0.8 shoulder injuries per 1,000 AE's (Table 1). Median days lost to an arm injury was 58 (11, 106).

ARM INJURY BURDEN PREDICTION MODEL

Model development following tenfold internal validation demonstrated a RMSE of 11.9 days, 0.80 R^2, and a calibration slope of 0.98 (95% CI: 0.92, 1.04); Figure 1). Full model equation is reported in Table 2.

The zero-inflated negative binomial model incorporates two models (count and zero) into one comprehensive model. As such, the count and zero model should be considered one model

ELBOW INJURY BURDEN PREDICTION MODEL

Elbow injury burden model development demonstrated a RMSE of 21.3 days, 0.42 R^2, and a calibration slope of 1.25 (95% CI: 1.16, 1.34; Figure 2).

SHOULDER INJURY BURDEN PREDICTION MODEL

Shoulder injury burden model development demonstrated a RMSE of 17.9 days, 0.57 R^2, and a calibration slope of 1.01 (95% CI: 0.92, 1.10; Figure 3).

SENSITIVITY ANALYSIS

The inclusion of only modifiable predictors within the prediction model decreased performance compared to the original model (RMSE: 21.2, R^2: 0.42, Calibration: 1.12 (95% CI: 0.99, 1.25).

DISCUSSION

SUMMARY

This prediction model demonstrated excellent performance, as demonstrated by the high R^2 and calibration slope. Due to the model error of 12 days, predictions of one to 14 days should be interpreted with caution. Stratifying by predicting individual elbow or shoulder burden decreased prediction model performance. Including only modifiable predictors demonstrated decreased prediction performance.

CLINICAL IMPLICATIONS

Clinicians integrate a plethora of skills, tools, and experience to keep athletes on the field.[55] The complexity associated with examination and performance, including advances in technology,[56] load monitoring,[57] and rehabilitation and performance testing,[23] necessitates the need employ these tools and information. Within baseball, athletes are physically examined in spring training and traditionally this information is used to assess each individual risk factor.[9] This prediction model incorporated multiple predictors into one cohesive model to calculate a predicted number of days lost to arm injury, ranging from 0 to 90. These individual injury burden predictions are meant to help reduce complexity of a difficult issue through clarifying prognosis,[58] to improve a clinician's ability to care for their patients.

This model reported an RMSE of 12 days. What this means clinically is that for any predicted days of injury burden, the actual number of days lost to injury will be within 12 days of the predicted value. While an error rate of 12 days for higher burden injuries (i.e., >30 days) may not alter clinical interpretation, a difference between 0 and 14 days could affect clinical decisions. To give an example, if the model predicted 3 days lost to injury, the days lost could be between one and 15 days. This would be clinically interpreted as a minor to moderate arm injury. On the other hand, if a pitcher was predicted to have 60 days lost to injury, a 12 day difference (48 to 72 days) would not affect clinical decisions. The pitcher would still be identified as a pitcher who could sustain a serious arm injury during the season. These findings suggest that this injury burden model can help identify pitchers who are at risk for sustaining a high injury burden during the season. However, these models should be interpreted with caution for predicted injury burden for one to 14 days.

Table 1. Characteristics of the pitchers included in the study. Values are percentages unless stated otherwise.

	All Pitchers (n = 297)	Non-Arm Injured Pitchers (n = 203)	Arm Injured Pitchers (n = 94)	Elbow Injury (n = 40)	Shoulder Injury (n = 64)
Mean (SD) Age (years)	23.0 (2.2)	22.8 (2.2)	23.2 (2.0)	23.2 (2.4)	22.8 (2.1)
Hand Dominance					
Left	14%	19%	22%	16%	16%
Right	86%	81%	78%	84%	84%
Mean (SD) BMI (kg/m2)	24.8 (2.2)	24.6 (2.3)	24.8 (2.1)	25.1 (2.3)	24.9 (2.3)
Pitching Role					
Starting Pitcher	55%	52%	62%	52%	58%
Relief Pitcher	45%	48%	38%	48%	42%
Continent of Origin					
North America	71%	66%	83%	68%	80%
Latin America	29%	34%	17%	32%	20%
Years within Professional Baseball					
1-2 years	73%	80%	57%	42%	58%
3-4 years	22%	3%	33%	46%	32%
5+ years	5%	1%	11%	12%	10%
Mean (SD) Seasonal Pitch Load	900 (657)	991 (643)	548 (615)	468 (594)	572 (609)
Mean (SD) Dominant Total Range of Motion in degrees	160 (13)	159 (13.2)	159 (11.8)	162 (12)	159 (14)
Mean (SD) Dominant Horizontal Adduction in degrees	-2.3 (12.2)	-2.8 (11.9)	-1.1 (14.6)	-1.4 (12.6)	-0.2 (13.2)
Mean (SD) Humeral Torsion Difference in degrees	18.2 (13.7)	18.5 (13.5)	16.2 (11.9)	17.3 (13.4)	15.9 (12.8)
Kinematic Chain Injury	15%	3%	11%	5%	13%

Total range of motion = Sum of shoulder external and internal rotation range of motion
Humeral torsion difference is calculated as dominant − nondominant shoulder

INDIVIDUAL ELBOW AND SHOULDER MODELS

Elbow and shoulder injury burden models demonstrated decreased performance compared to the primary model and had reduced sample sizes therefore, should be interpreted with caution. The discrepancies between elbow and shoulder injury burden may decrease the performance of these models. Elbow and shoulder injuries can have different risk factors,[9,10,59] with mechanisms of injury occurring during different points within the pitching motion.[19,60] Different clinical interventions may be required to reduce shoulder verses elbow injury risk.[61,62] The contrasting injury burden risk factors between shoulders and elbows may decrease the prediction precision of individual shoulder or elbow models. Future research is required to develop separate shoulder and elbow burden prediction models.

INCLUSION OF ONLY MODIFIABLE PREDICTORS

Including only modifiable predictors demonstrated decreased injury burden prediction compared to the inclusion of modifiable and non-modifiable predictors. Previous work[14] has called for the inclusion of both modifiable and non-modifiable predictors to improve understanding and ultimately decrease sports injuries. However, while this has been proposed for many years,[14] the majority of current ev-idence only includes isolated modifiable factors when assessing injury risk.[9] As these results suggest, without the inclusion of non-modifiable predictors, clinicians are inhibited from having an improved understanding of current injury risk. While it may seem from a clinical perspective that including non-modifiable predictors does not improve how a clinician will treat a patient, a more comprehensive injury risk examination allows the clinician to better identify athletes at risk for sustaining an injury.[63] These prediction models are not recommended to be used to identify which specific interventions or factors should be the focus of care.[28,29] Prediction models are not causal, as such, if a particular predictor is "significant" with a specific athlete, this does not mean that particular predictor should be intervened upon. A thorough physical exam and clinical reasoning should be used to identify which tests and measures to intervene on.

STRENGTHS AND POTENTIAL LIMITATIONS

This study utilized a seven-year prospective cohort that was specifically designed to assess arm injuries in professional baseball players. All data collectors were physical therapists, specifically trained and evaluated for reliability in the physical examination techniques, improving the consistency and reliability of these data. Only a small proportion
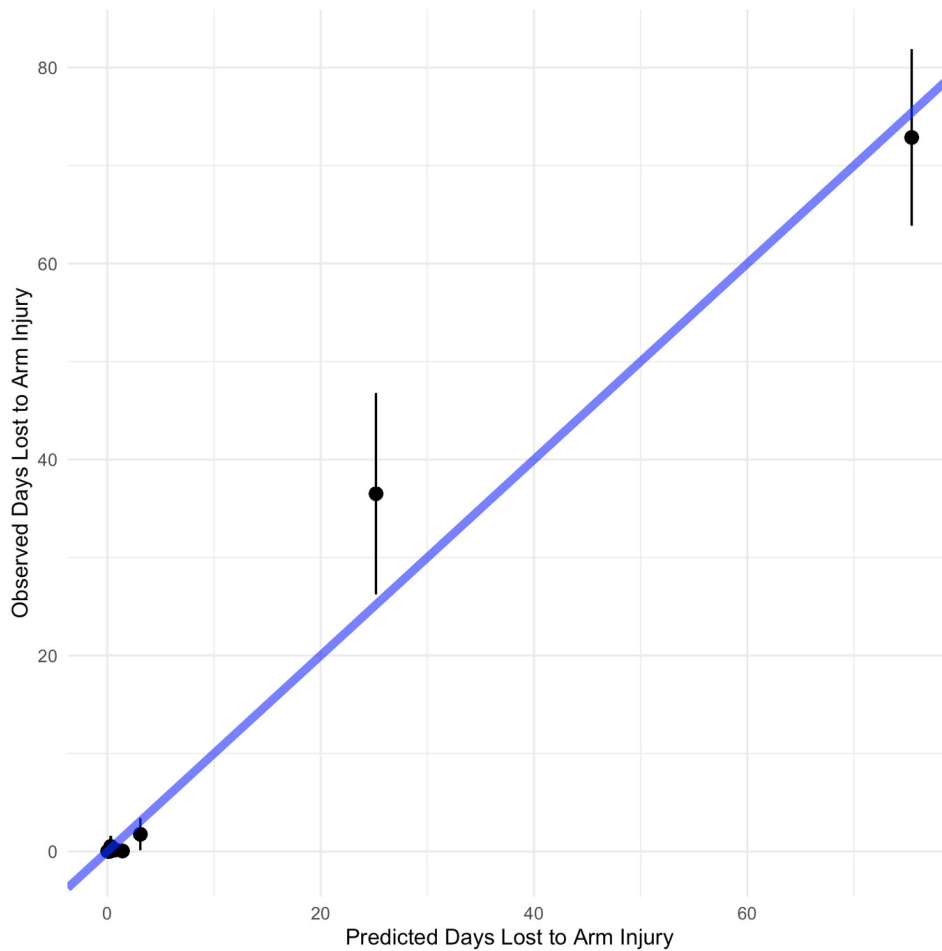
Figure 1. Calibration Plot of the Arm Injury Burden Prediction Model.

Calibration is the relationship between predicted and actual probability of the event. Perfect calibration would be a 45-degree line. A calibration line below the blue line is over-predicting risk. A calibration line above the blue line is under-predicting risk.

of data were missing, allowing for a complete case analysis, increasing the validity of these findings. An *a priori* sample size calculation was performed to create a stable prediction model, increasing the utility of these findings. Internal validation was performed in order to shrink optimism, increasing the generalizability of these results. The full model is reported, increasing transparency and usefulness for future validation.[64] Time lost to injury was truncated at 90 days. Pitchers may be placed on the injured list longer than 90 days; however, truncation was performed at 90 days due to the paucity of players sustaining time loss greater than this time period and the organizational significance of this time period. A small proportion of pitchers were excluded due to participating MLB spring training during data collection. These pitchers were predominantly at the AAA or AA level, decreasing the generalizability of these results to high MiLB players. Injury history was limited to professional baseball seasons and orthopaedic surgery prior to signing a professional baseball contract. As pitchers may sustain arm injuries that last only a few days or a week, there is the potential for residual confounding. All impairment data were collected prior to the season. As predictors can change throughout the season, this decreases the clinical utility of these results. Player salary and signing

bonuses were not included in these analyses. As MLB organizational investment may influence time loss to injury, this decreases the precision in these models. This model was not externally validated. It is recommended that external validation should be performed prior to integrating in a clinical setting. This prediction model should be used with caution without further validation.

CONCLUSIONS

This professional baseball injury burden model may have clinical utility in predicting pitchers that are at risk for sustaining a high injury burden within the season. However, caution should be advised with predictions between 1 to 14 days lost to arm injury as this is below the prescribed model error. The inclusion of modifiable and non-modifiable predictors demonstrated improved prediction performance, suggesting that prediction models should include both types of predictors when evaluating injury risk. Separate elbow and shoulder prediction models demonstrated decreased performance and should be interpreted with caution due to low sample size. Further research is required to externally validate this model to understand the generalizability of these findings.

Table 2. Arm Injury Burden Prediction Model.

	Coefficient	*95% CI
Count Portion of the Model		
Intercept	1.62	0.11, 210
Age	-0.02	0.89, 1.05
Body Mass Index	0.05	0.99, 1.16
Right Arm Dominance ø	0.04	0.68, 1.72
Starting Pitcher¥	0.08	0.76, 1.38
Number of Pitching Appearances	-0.02	0.97, 0.99
Dominant Shoulder Total Range of Motion	0.01	1.00, 1.03
Dominant Shoulder Horizontal Adduction	-0.01	0.98, 0.99
Previous Arm Injury History	0.07	0.61, 1.47
2 to 4 Years Played Professionally‡	-0.37	0.45, 0.83
5+ Years Played Professionally‡	0.08	0.60, 2.03
Received Individualized Injury Prevention Program	-0.11	0.55, 1.01
Continent of Origin§	0.06	0.76, 1.08
Exposure Days	-0.006	0.98, 0.99
Zero Portion of the Model		
Intercept	-5.82	0.00, 8.99
Right Arm Dominanceø	-0.21	0.01, 1.87
Starting Pitcher¥	-0.06	0.15, 1.78
Number of Pitching Appearances	0.01	0.97, 1.08
Kinematic Chain Injury	0.86	2.12, 73
Continent of Origin§	0.07	0.74, 11.16
Exposure Days	0.05	1.03, 1.07

The zero-inflated negative binomial model incorporates two models (count and zero) into one comprehensive model. As such, the count and zero model should be considered one model

95% CI = 95% Confidence Interval

*95% confidence intervals are exponentially transformed

øReference is Left

¥Reference is relief pitcher

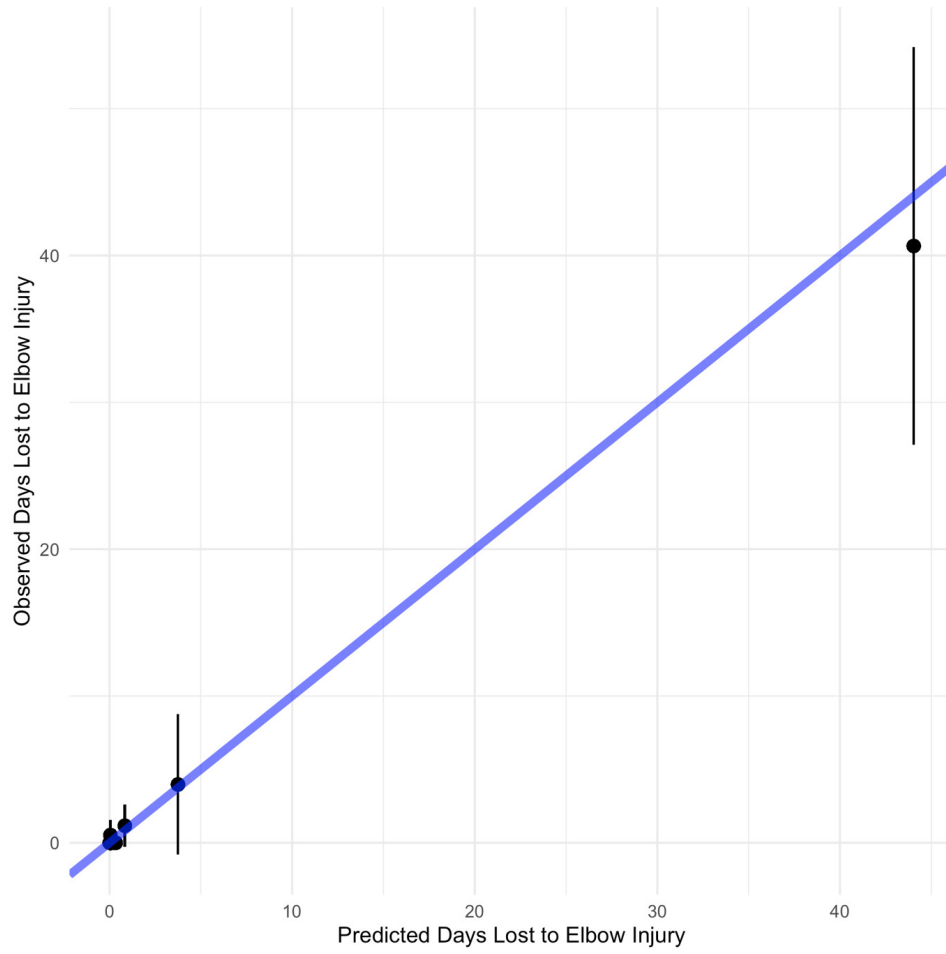‡Reference is 0 to 1 years played professionally

§Reference is North America

CONFLICTS OF INTEREST

All authors declare no conflicts of interest.

Figure 2. Calibration Plot of the Elbow Injury Burden Prediction Model.

Calibration is the relationship between predicted and actual probability of the event. Perfect calibration would be a 45-degree line. A calibration line below the blue line is over-predicting risk. A calibration line above the blue line is under-predicting risk.
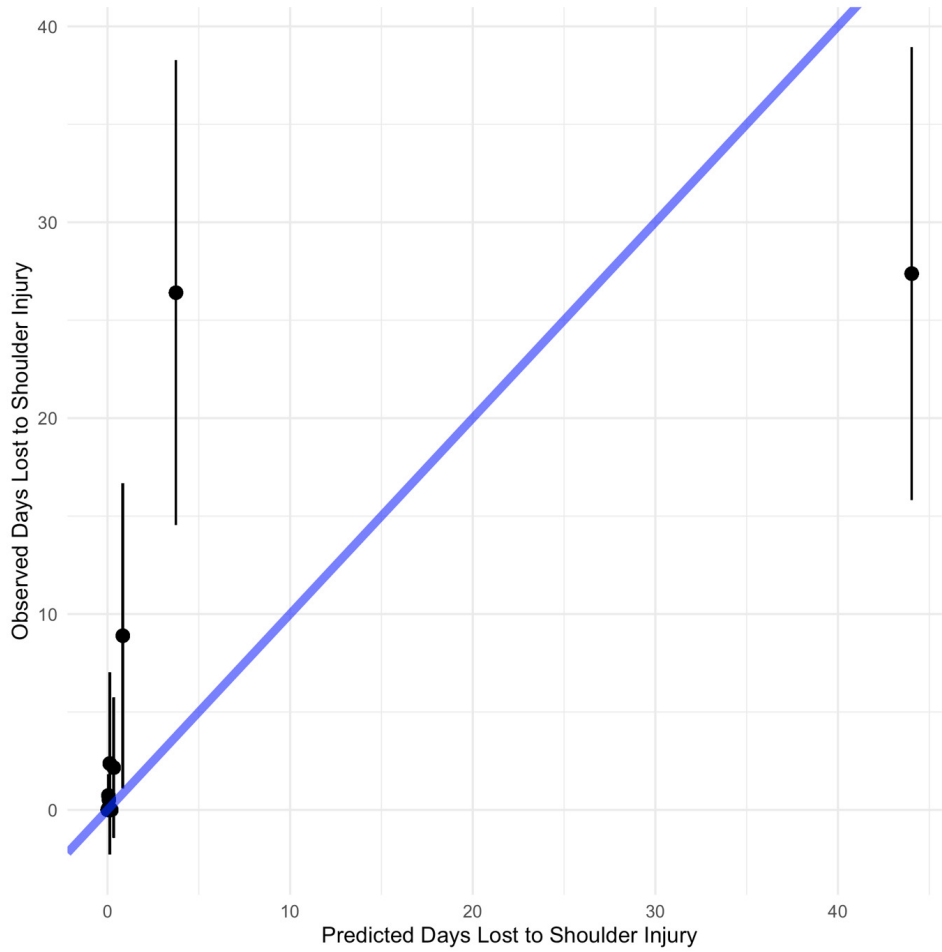
Figure 3. Calibration Plot of the Shoulder Injury Burden Prediction Model.

Calibration is the relationship between predicted and actual probability of the event. Perfect calibration would be a 45-degree line. A calibration line below the blue line is over-predicting risk. A calibration line above the blue line is under-predicting risk.

REFERENCES

1. Camp CL, Conte S, D'Angelo J, Fealy SA. Epidemiology of ulnar collateral ligament reconstruction in Major and Minor League Baseball pitchers: comprehensive report of 1429 cases. *J Shoulder Elbow Surg.* 2018;27(5):871-878. doi:10.1016/j.jse.2018.01.024

2. Conte S, Requa RK, Garrick JG. Disability days in major league baseball. *Am J Sport Med.* 2001;29(4):431-436. doi:10.1177/03635465010290040801

3. Camp CL, Dines JS, van der List JP, et al. Summative report on time out of play for Major and Minor League Baseball: an analysis of 49,955 injuries from 2011 through 2016. *Am J Sport Med.* 2018;46(7):1727-1732. doi:10.1177/0363546518765158

4. Conte SA, Fleisig GS, Dines JS, et al. Prevalence of ulnar collateral ligament surgery in professional baseball players. *Am J Sport Med.* 2015;43(7):1764-1769. doi:10.1177/0363546515580792

5. Bullock GS, Uhan J, Harriss EK, Arden NK, Filbay SR. The Relationship Between Baseball Participation and Health: A Systematic Scoping Review. *J Orthop Sport Phys Ther.* 2020;50(2):55-66. doi:10.2519/jospt.2020.9281

6. Knowles SB, Marshall SW, Miller T, et al. Cost of injuries from a prospective cohort study of North Carolina high school athletes. *Inj Prevent.* 2007;13(6):416-421. doi:10.1136/ip.2006.014720

7. Conte S, Camp CL, Dines JS. Injury Trends in Major League Baseball Over 18 Seasons: 1998-2015. *Am J Orthop (Belle Mead NJ).* 2016;45(3):116-123.

8. Bahr R, Clarsen B, Ekstrand J. Why we should focus on the burden of injuries and illnesses, not just their incidence. *Br J Sport Med.* 2018;52(16):1018-1021. doi:10.1136/bjsports-2017-098160

9. Wilk KE, Macrina LC, Fleisig GS, et al. Deficits in glenohumeral passive range of motion increase risk of elbow injury in professional baseball pitchers: a prospective study. *Am J Sport Med.* 2014;42(9):2075-2081. doi:10.1177/0363546514538391

10. Noonan TJ, Thigpen CA, Bailey LB, et al. Humeral torsion as a risk factor for shoulder and elbow injury in professional baseball pitchers. *Am J Sport Med.* 2016;44(9):2214-2219. doi:10.1177/0363546516648438

11. Lyman S, Fleisig GS, Andrews JR, Osinski ED. Effect of pitch type, pitch count, and pitching mechanics on risk of elbow and shoulder pain in youth baseball pitchers. *Am J Sport Med.* 2002;30(4):463-468. doi:10.1177/03635465020300040201

12. Garner JC, MacDonald C, Wade C, Johnson A, Ford MA. The influence of body composition on youth throwing kinetics. *Ped Ex Sci.* 2011;23(3):379-387. doi:10.1123/pes.23.3.379

13. Karakolis T, Bhan S, Crotin RL. An inferential and descriptive statistical examination of the relationship between cumulative work metrics and injury in Major League Baseball pitchers. *J Strength Cond Res.* 2013;27(8):2113-2118. doi:10.1519/jsc.0b013e318278 5059

14. Meeuwisse WH. Assessing causation in sport injury: a multifactorial model. *Clin J Sport Med.* 1994;4(3):166-170. doi:10.1097/00042752-199407000-00004

15. Saper MG, Pierpoint LA, Liu W, Comstock RD, Polousky JD, Andrews JR. Epidemiology of shoulder and elbow injuries among United States high school baseball players: school years 2005-2006 through 2014-2015. *Am J Sport Med.* 2018;46(1):37-43. doi:10.1177/0363546517734172

16. Li X, Zhou H, Williams P, et al. The epidemiology of single season musculoskeletal injuries in professional baseball. *Orthop Rev.* 2013;5(1):3. doi:10.4081/or.2013.e3

17. Ciccotti MG, Pollack KM, Ciccotti MC, et al. Elbow injuries in professional baseball: epidemiological findings from the Major League Baseball Injury Surveillance System. *Am J Sport Med.* 2017;45(10):2319-2328. doi:10.1177/0363546517706964

18. Philp F, Al-Shallawi A, Kyriacou T, Blana D, Pandyan A. Improving predictor selection for injury modelling methods in male footballers. *BMJ Open Sport Exerc Med.* 2020;6(1):e000634. doi:10.1136/bmjsem-2019-000634

19. Fleisig GS, Andrews JR, Dillman CJ, Escamilla RF. Kinetics of baseball pitching with implications about injury mechanisms. *Am J Sport Med.* 1995;23(2):233-239. doi:10.1177/036354659502300218

20. Erikkson E. Sports injuries of the knee ligaments: their diagnosis, treatment, rehabilitation, and prevention. *Med Sci Sport*. 1976;8(3):133-144. [doi:10.1249/00005768-197600830-00001](doi:10.1249/00005768-197600830-00001)

21. Cleland J, Koppenhaver S, Su J. *Netter's Orthopaedic Clinical Examination: An Evidence-Based Approach*. Elsevier Health Sciences; 2015.

22. Bahr R. Why screening tests to predict injury do not work—and probably never will...: a critical review. *Br J Sport Med*. 2016;50(13):776-780. doi:10.1136/bjsports-2016-096256

23. Hughes T, Sergeant JC, van der Windt DA, Riley R, Callaghan MJ. Periodic health examination and injury prediction in professional football (Soccer): theoretically, the prognosis is good. *Sport Med*. 2018;48(11):2443-2448. [doi:10.1007/s40279-018-0928-y](doi:10.1007/s40279-018-0928-y)

24. Bittencourt NFN, Meeuwisse WH, Mendonça LD, Nettel-Aguirre A, Ocarino JM, Fonseca ST. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. *Br J Sport Med*. 2016;50(21):1309-1314. doi:10.1136/bjsports-2015-095850

25. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. *Int J Epidemiol*. 2010;39(1):97-106. [doi:10.1093/ije/dyp296](doi:10.1093/ije/dyp296)

26. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338(feb23 1):b375-b375. [doi:10.1136/bmj.b375](doi:10.1136/bmj.b375)

27. Moons KGM, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart*. 2012;98(9):683-690. doi:10.1136/heartjnl-2011-301246

28. Bullock GS, Hughes T, Sergeant JC, Callaghan MJ, Riley R, Collins G. Methods matter: clinical prediction models will benefit sports medicine practice, but only if they are properly developed and validated. *Br J Sport Med*. 2021;55(23):1319-1321. doi:10.1136/bjsports-2021-104329

29. Bullock GS, Hughes T, Sergeant JC, Callaghan MJ, Riley RD, Collins GS. Clinical prediction models in sports medicine: A guide for clinicians and researchers. *J Orthop Sport Phys Ther*. 2021;51(10):517-525. doi:10.2519/jospt.2021.10697

30. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. *Am J Cardiol*. 1976;38(1):46-51. [doi:10.1016/0002-9149(76)90061-8](doi:10.1016/0002-9149(76)90061-8)

31. Quatman CE, Quatman CC, Hewett TE. Prediction and prevention of musculoskeletal injury: a paradigm shift in methodology. *Br J Sport Med*. 2009;43(14):1100-1107. doi:10.1136/bjsm.2009.065482

32. Van Calster B, Vickers AJ. Calibration of risk prediction models: impact on decision-analytic performance. *Med Decis Making*. 2015;35(2):162-169. [doi:10.1177/0272989x14547233](doi:10.1177/0272989x14547233)

33. Shanley E, Kissenberth MJ, Thigpen CA, et al. Preseason shoulder range of motion screening as a predictor of injury among youth and adolescent baseball pitchers. *J Shoulder Elbow Surg*. 2015;24(7):1005-1013. [doi:10.1016/j.jse.2015.03.012](doi:10.1016/j.jse.2015.03.012)

34. Lintner D, Mayol M, Uzodinma O, Jones R, Labossiere D. Glenohumeral internal rotation deficits in professional pitchers enrolled in an internal rotation stretching program. *Am J Sport Med*. 2007;35(4):617-621. [doi:10.1177/0363546506296736](doi:10.1177/0363546506296736)

35. Meister K, Day T, Horodyski M, Kaminski TW, Wasik MP, Tillman S. Rotational motion changes in the glenohumeral joint of the adolescent/Little League baseball player. *Am J Sport Med*. 2005;33(5):693-698. [doi:10.1177/0363546504269936](doi:10.1177/0363546504269936)

36. Laudner KG, Stanek JM, Meister K. Assessing posterior shoulder contracture: the reliability and validity of measuring glenohumeral joint horizontal adduction. *J Athl Train*. 2006;41(4):375-380.

37. Myers JB, Oyama S, Wassinger CA, et al. Reliability, precision, accuracy, and validity of posterior shoulder tightness assessment in overhead athletes. *Am J Sport Med*. 2007;35(11):1922-1930. [doi:10.1177/0363546507304142](doi:10.1177/0363546507304142)

38. Shanley E, Rauh MJ, Michener LA, Ellenbecker TS, Garrison JC, Thigpen CA. Shoulder range of motion measures as risk factors for shoulder and elbow injuries in high school softball and baseball players. *Am J Sport Med*. 2011;39(9):1997-2006. [doi:10.1177/0363546511408876](doi:10.1177/0363546511408876)

39. Shanley E, Bailey L, Sandago MP, Pinkerton A, Singleton SB, Thigpen CA. The use of a pitch count estimator to calculate exposure in collegiate baseball pitchers. *Phys Ther Sport*. 2015;16(4):344-348. doi:10.1016/j.ptsp.2015.02.002

40. Rae K, Orchard J. The Orchard sports injury classification system (OSICS) version 10. *Clin J Sport Med*. 2007;17(3):201-204. doi:10.1097/jsm.0b013e318059b536

41. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. doi:10.1136/bmj.m441

42. Simning A, Orth J, Temkin-Greener H, Li Y. Patients discharged from higher-quality skilled nursing facilities spend more days at home. *Health Services Res*. 2021;56(1):102-111. doi:10.1111/1475-6773.13543

43. Wilk KE, Macrina LC, Fleisig GS, et al. Correlation of glenohumeral internal rotation deficit and total rotational motion to shoulder injuries in professional baseball pitchers. *Am J Sport Med*. 2011;39(2):329-335. doi:10.1177/0363546510384223

44. Reinold MM, Wilk KE, Macrina LC, et al. Changes in shoulder and elbow passive range of motion after pitching in professional baseball players. *Am J Sport Med*. 2008;36(3):523-527. doi:10.1177/0363546507308935

45. Noonan TJ, Thigpen CA, Bailey LB, et al. Humeral Torsion as a Risk Factor for Shoulder and Elbow Injury in Professional Baseball Pitchers. *Am J Sport Med*. 2016;44(9):2214-2219. doi:10.1177/0363546516648438

46. Helmkamp JK, Bullock GS, Rao A, Shanley E, Thigpen C, Garrigues GE. The Relationship Between Humeral Torsion and Arm Injury in Baseball Players: A Systematic Review and Meta-analysis. *Sports Health*. 2020;12(2):132-138. doi:10.1177/1941738119900799

47. Polster JM, Bullen J, Obuchowski NA, Bryan JA, Soloff L, Schickendantz MS. Relationship between humeral torsion and injury in professional baseball pitchers. *Am J Sport Med*. 2013;41(9):2015-2021. doi:10.1177/0363546513493249

48. Takeuchi S, Yoshida M, Sugimoto K, Tsuchiya A, Takenaga T, Goto H. The differences of humeral torsion angle and the glenohumeral rotation angles between young right-handed and left-handed pitchers. *J Shoulder Elbow Surg*. 2019;28(4):678-684. doi:10.1016/j.jse.2018.09.002

49. Sciascia A, Thigpen C, Namdari S, Baldwin K. Kinetic chain abnormalities in the athletic shoulder. *Sports Med Arthrosc*. 2012;20(1):16-21. doi:10.1097/jsa.0b013e31823a021f

50. Bullock GS, Thigpen CA, Noonan TK, Kissenberth MJ, Shanley E. Initial Kinematic Chain Injuries Increase Hazard of Subsequent Arm Injuries in Professional Baseball Pitchers. *J Shoulder Elbow Surg*. 2022;31(9):1773-1781. doi:10.1016/j.jse.2022.04.016

51. Bullock GS, Thigpen CA, Collins GS, et al. Starting pitchers demonstrate a greater hazard of arm injury compared to relief pitchers. *J Athl Train*. 2022;57(1):65-71. doi:10.4085/1062-6050-0262.21

52. Bradbury JC, Forman SL. The impact of pitch counts and days of rest on performance among major-league baseball pitchers. *J Strength Cond Res*. 2012;26(5):1181-1187. doi:10.1519/jsc.0b013e31824e16fe

53. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Statistical Soc B*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x

54. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594. doi:10.1136/bmj.g7594

55. Levy AS, Lintner S, Kenter K, Speer KP. Intra-and interobserver reproducibility of the shoulder laxity examination. *Am J Sport Med*. 1999;27(4):460-463. doi:10.1177/03635465990270040901

56. Osborne B. Legal and Ethical Implications of Athletes' Biometric Data Collection in Professional Sport. *Marq Sports L Rev*. 2017;28:37.

57. Gabbett TJ, Jenkins DG. Relationship between training load and injury in professional rugby league players. *J Sci med Sport*. 2011;14(3):204-209. doi:10.1016/j.jsams.2010.12.002

58. Wynants L, Collins G, Van Calster B. Key steps and common pitfalls in developing and validating risk models. *BJOG*. 2017;124(3):423-432. doi:10.1111/1471-0528.14170

59. Wilk KE, Macrina LC, Fleisig GS, et al. Deficits in glenohumeral passive range of motion increase risk of shoulder injury in professional baseball pitchers: a prospective study. *Am J Sport Med*. 2015;43(10):2379-2385. doi:10.1177/0363546515594380

60. Werner SL, Guido JA Jr, Stewart GW, McNeice RP, VanDyke T, Jones DG. Relationships between throwing mechanics and shoulder distraction in collegiate baseball pitchers. *J Shoulder Elbow Surg*. 2007;16(1):37-42. doi:10.1016/j.jse.2006.05.007

61. Reinold MM, Gill TJ, Wilk KE, Andrews JR. Current concepts in the evaluation and treatment of the shoulder in overhead throwing athletes, part 2: injury prevention and treatment. *Sports Health*. 2010;2(2):101-115. doi:10.1177/1941738110362518

62. Wilk KE, Reinold MM, Andrews JR. Rehabilitation of the thrower's elbow. *Clin Sport Med*. 2004;23(4):765-801. doi:10.1016/j.csm.2004.06.006

63. Shanley E, Thigpen CA, Collins GS, et al. Including modifiable and non-modifiable factors improves injury risk assessment in professional baseball pitchers. *J Orthop Sport Phys Ther*. 2022(0):1-42.

64. Bullock GS, Hughes T, Arundale AAJH, Ward P, Collins GS, Kluzek S. Black Box Prediction Methods in Sports Medicine Deserve a Red Card for Reckless Practice: A Change of Tactics is Needed to Advance Athlete Care. *Sports Med*. Published online October 14, 2022:1-7. doi:10.1007/s40279-022-01771-3

SUPPLEMENTARY MATERIALS

Appendix 4

Download: https://ijspt.scholasticahq.com/article/39741-development-of-an-injury-burden-prediction-model-in-professional-baseball-pitchers/attachment/103300.docx

Appendix 3

Download: https://ijspt.scholasticahq.com/article/39741-development-of-an-injury-burden-prediction-model-in-professional-baseball-pitchers/attachment/103301.docx

Appendix 2

Download: https://ijspt.scholasticahq.com/article/39741-development-of-an-injury-burden-prediction-model-in-professional-baseball-pitchers/attachment/103302.docx

Appendix 1

Download: https://ijspt.scholasticahq.com/article/39741-development-of-an-injury-burden-prediction-model-in-professional-baseball-pitchers/attachment/103303.docx