Predictive Modeling of Hamstring Strain Injuries in Elite Australian Footballers

JOSHUA D. RUDDY[1], ANTHONY J. SHIELD[2], NIRAV MANIAR[1], MORGAN D. WILLIAMS[3], STEVEN DUHIG[4], RYAN G. TIMMINS[1], JACK HICKEY[1], MATTHEW N. BOURNE[4], and DAVID A. OPAR[1]

[1]School of Exercise Science, Australian Catholic University, Melbourne, AUSTRALIA; [2]School of Exercise and Nutrition Science, Faculty of Health, Queensland University of Technology, Brisbane, AUSTRALIA; [3]School of Health, Sport and Professional Practice, Faculty of Life Sciences and Education, University of South Wales, UNITED KINGDOM; and [4]School of Allied Health Sciences, Griffith University, Gold Coast, AUSTRALIA

ABSTRACT

RUDDY, J. D., A. J. SHIELD, N. MANIAR, M. D. WILLIAMS, S. DUHIG, R. G. TIMMINS, J. HICKEY, M. N. BOURNE, and D. A. OPAR. Predictive Modeling of Hamstring Strain Injuries in Elite Australian Footballers. *Med. Sci. Sports Exerc.*, Vol. 50, No. 5, pp. 906–914, 2018. **Purpose**: Three of the most commonly identified hamstring strain injury (HSI) risk factors are age, previous HSI, and low levels of eccentric hamstring strength. However, no study has investigated the ability of these risk factors to predict the incidence of HSI in elite Australian footballers. Accordingly, the purpose of this prospective cohort study was to investigate the predictive ability of HSI risk factors using machine learning techniques. **Methods**: Eccentric hamstring strength, demographic and injury history data were collected at the start of preseason for 186 and 176 elite Australian footballers in 2013 and 2015, respectively. Any prospectively occurring HSI were reported to the research team. Using various machine learning techniques, predictive models were built for 2013 and 2015 within-year HSI prediction and between-year HSI prediction (2013 to 2015). The calculated probabilities of HSI were compared with the injury outcomes and area under the curve (AUC) was determined and used to assess the predictive performance of each model. **Results**: The minimum, maximum, and median AUC values for the 2013 models were 0.26, 0.91, and 0.58, respectively. For the 2015 models, the minimum, maximum and median AUC values were, correspondingly, 0.24, 0.92, and 0.57. For the between-year predictive models the minimum, maximum, and median AUC values were 0.37, 0.73, and 0.52, respectively. **Conclusions**: Although some iterations of the models achieved near perfect prediction, the large ranges in AUC highlight the fragility of the data. The 2013 models performed slightly better than the 2015 models. The predictive performance of between-year HSI models was poor however. In conclusion, risk factor data cannot be used to identify athletes at an increased risk of HSI with any consistency. **Key Words**: INJURY PREDICTION, MACHINE LEARNING, ECCENTRIC HAMSTRING STRENGTH, HAMSTRING INJURY RISK

Hamstring strain injuries (HSI) are the most common injury in elite Australian football (1) and can have significant physical (2–4) and financial consequences for an athlete and their sporting organization (5). As such, risk factors for HSI have received much attention (6–12). The most common risk factors identified in Australian footballers are a history of HSI and increasing age (9). More recently, work has focused on factors that can be targeted through intervention to potentially reduce the risk of HSI. Previous work reported that elite Australia footballers with levels of eccentric hamstring strength (assessed during the Nordic hamstring exercise) below 256 N at the start of preseason were 2.7 times more likely to sustain a HSI that season (10). However, this cutoff was determined retrospectively from the data it was applied to, and although deductive methodologies such as this are able to establish a link between a factor such as eccentric hamstring strength and injury risk, they cannot be used to predict injury (13). Recent commentary has highlighted the need to estimate the predictive ability of risk factors by applying a predetermined cutoff to different samples (13). Injuries, such as HSI, however, are typically the result of the interactions between multiple factors, and univariate cutoff values are likely to display poor predictive capacity (14,15).

Although multivariate relationships between HSI risk factors have been examined previously, these investigations typically use a linear approach to identifying the interactions between these variables (15,16). These interactions, however, are more likely to be nonlinear (14–16). It has been suggested that moving away from the identification of univariate risk factor cutoffs and instead toward injury pattern recognition will improve the ability to predict and ultimately prevent injury (14). Supervised learning techniques, which are a type of machine learning, have been proposed as

appropriate methodologies to account for the complex, nonlinear interactions between risk factors and to recognize patterns that lead to injury (14). Supervised learning is the process by which a data set with a known outcome variable (i.e., injured or uninjured), referred to as training data, is used to identify patterns and predict the unknown outcome variable of an independent data set, referred to as testing data (17). No study has developed supervised learning models, using data from multiple teams measured across multiple seasons, to predict the occurrence of HSI. Studies identifying risk factors and their associated cutoffs can provide important information regarding the etiology of injuries (18). Research of this nature can ultimately help to inform injury prevention practices and guide risk mitigation strategies (19). However, investigations using supervised learning techniques are needed to determine the predictive ability of previously reported HSI risk factors such as eccentric hamstring strength, age, and previous HSI.

Therefore, the aim of the current study was to investigate whether supervised learning techniques could be used to predict the occurrence of HSI, using risk factor and demographic data measured across multiple seasons as predictor variables.

METHODS

Study design. Data for this prospective cohort study were collected during the 2013 and 2015 Australian Football League (AFL) seasons. The 2013 study period was conducted from November 2012 to September 2013, and the 2015 study period was conducted from November 2014 to September 2015. The authors would like to note that data collected during the 2013 study period has been previously published (10). Demographic and lower-limb injury history data for each athlete were provided to the research team at the beginning of the 2013 and 2015 study periods. Eccentric hamstring strength was assessed at start of preseason training for 2013 (November 2012) and 2015 (November 2014). Throughout the study periods, any prospectively occurring HSI were reported to the research team. This study was approved by the Queensland University of Technology Human Research Ethics Committee (approval number: 1100001116).

Participants. Seven teams (39% of the total competition) competing in the AFL participated in at least one of the study periods. Three of these teams participated in both the 2013 and 2015 study periods. Four teams only participated in either the 2013 or 2015 study period. Each athlete was provided with a plain language statement outlining the study and provided informed written consent. In total, 186 and 176 elite Australian footballers agreed to participate in the 2013 and 2015 study periods, respectively. Although data for 210 athletes were previously reported (10), eccentric hamstring strength was only assessed at the start of preseason for 186 athletes during the 2013 study period. Due to deidentification of the data, it was not possible to determine which athletes provided data for both study periods.

Demographic and injury history data. Demographic data for each athlete were provided to the research team at the beginning of each study period. This included age (yr), stature (cm), mass (kg), and primary playing position (forward, back, midfield, or ruck (11)). The medical staff for each of the participating teams also completed a questionnaire detailing the lower limb injury history of each athlete prior to the 2013 and 2015 seasons. This included history of HSI within the preceding 12 months and history of anterior cruciate ligament (ACL) injury at any stage during the athlete's career.

Eccentric hamstring strength assessment. The assessment of eccentric knee flexor strength during the Nordic hamstring exercise has been previously reported (10,20,21). The participants knelt on a padded board with their ankles secured superior to the lateral malleolus by ankle straps. These straps were attached to custom-made uniaxial load cells (Delphi Force Measurement, Gold Coast, Australia) with wireless data acquisition capabilities (Mantracourt, Devon, United Kingdom). After a warm-up set of three submaximal repetitions, participants were asked to perform three maximal repetitions of the Nordic hamstring exercise. The participants were instructed to lean forwards as slowly as possible whilst maximally resisting this movement with both limbs, keeping the trunk and hips neutral throughout. The athletes were encouraged to perform maximally on every repetition. The assessment took place within a week of each individual athlete's return to preseason training for the 2013 and 2015 seasons. Data were transferred to a computer at a 100-Hz through a wireless USB base station (Mantracourt, Devon, United Kingdom). The peak forces (measured in N) from each limb and repetition were averaged to determine athletes' eccentric hamstring strength. The absolute difference in the average force (N) between limbs was determined to be the between-limb imbalance.

Reporting of prospective HSI. A prospectively occurring HSI was defined as acute pain in the posterior thigh that resulted in disruption of the hamstring fibers, as confirmed by magnetic resonance imagining (MRI) (10). For all injuries that fulfilled these criteria, the relevant team doctor or physiotherapist completed a standard injury report form which detailed the limb that was injured, the location of the injury (i.e., proximal muscle-tendon junction, mid muscle belly etc.), mechanism of injury (i.e., high-speed running, jumping, and so on), severity of the injury determined from MRI, and the number of days taken to return to full training (10).

Modeling approaches. Two approaches were taken when selecting variables to include as input data for the predictive models. The first approach only included eccentric hamstring strength, age and previous HSI as predictors for future HSI. The second approach included eccentric hamstring strength, age, previous HSI, between-limb imbalance, previous ACL injury, stature, mass, and primary playing position. Both years (2013 and 2015) were first analyzed individually to predict within-year HSI. The modeling approach taken for

within-year HSI prediction can be found in Figure 1A. Between-year HSI prediction was then explored by allocating the 2013 data as training data and the 2015 data as testing data (Fig. 1B). To account for any variance caused by random sampling, data preprocessing and cross-validation, 10,000 iterations of each modeling approach were performed. All analyses were performed using the R statistical programming language (22).

Algorithms. Five different supervised learning techniques were used to build predictive models (see Table, Supplemental Digital Content 1, A brief description of each machine learning algorithm and all relevant parameters, http://links.lww.com/MSS/B163). These were:

- Naïve Bayes,
- Logistic regression,
- Random forest,
- Support vector machine,
- Neural network.

Explaining the mathematical functions of each of these models is beyond the scope of this article. However, naive Bayes and logistic regression algorithms were chosen for their simple probabilistic classification, whereas random forest, support vector machine, and neural network algorithms were chosen for their ability to model complex, nonlinear interactions amongst multiple predictor variables.

Data preprocessing. All continuous data (age, eccentric hamstring strength, stature, and mass) were standardized before building the predictive models. Standardization is common practice when implementing supervised learning techniques as some models may be sensitive to the vastly different ranges and magnitudes of predictor variables (17). This involves transforming the data so that the mean is equal to zero and the SD is equal to one. Data were standardized using the following equation:

$$x^1 = \frac{x - \bar{x}}{s}$$

where x equals the original value, x^1 equals the standardized value, $\bar{x}$ equals the mean and s equals the SD.

The input data for within-year HSI prediction were standardized to the relevant year (i.e., 2013 or 2015). The training data (2013) and the testing data (2015) for between-year HSI prediction were standardized independently. Another challenge when using statistical learning techniques, particularly to predict injury, is class imbalance (23). Although HSI is the most common injury in elite Australian football (1), uninjured athletes still outnumber athletes who sustain a HSI. In cases of class imbalance, supervised learning algorithms can achieve high accuracy by always predicting the more frequently occurring class (24). However, this high accuracy is only reflective of the underlying class distribution. Synthetic minority oversampling techniques (SMOTE) were developed to address the problem of having imbalanced classes (24). Synthetic minority oversampling techniques is a combination of undersampling and synthetic oversampling techniques (24). Undersampling is the process by which observations from the overrepresented class are
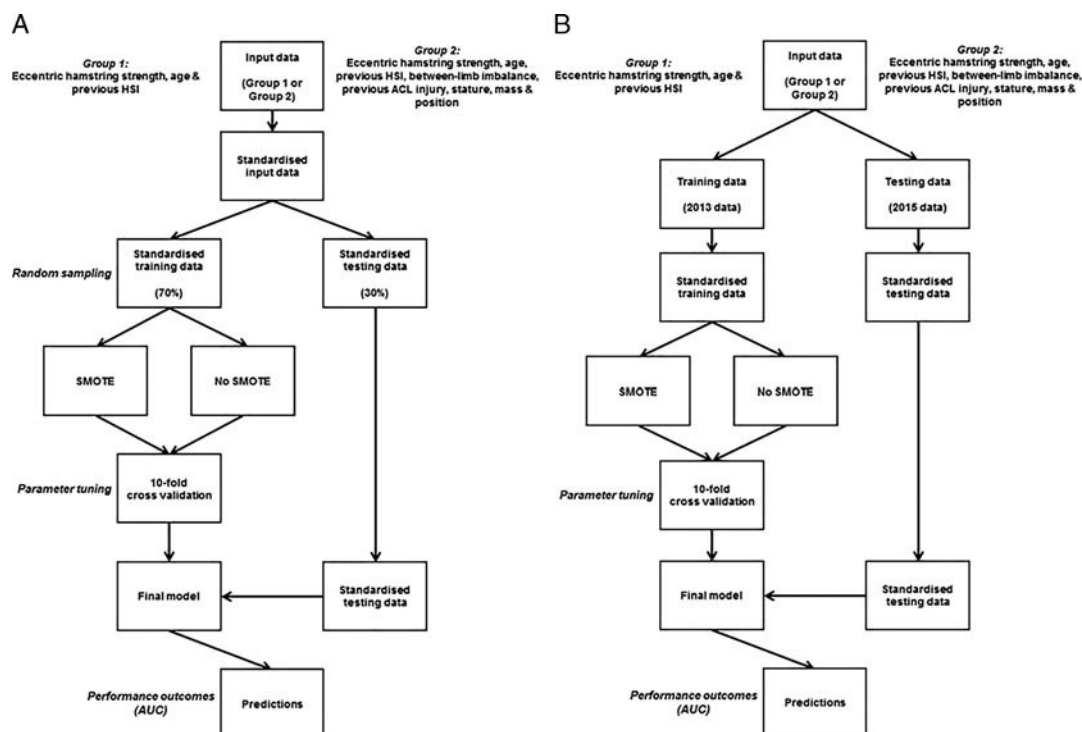

FIGURE 1—The modelling approaches for (A) within-year HSI prediction (used to analyse 2013 data and 2015 data independently), and (B) between-year HSI predictions (used to analyse 2013 data and 2015 data collectively). The performance outcome used was AUC. Ten thousand iterations of each modeling approach were performed.

randomly removed, whereas synthetic oversampling creates observations of the underrepresented class that have similar features to already existing observations (24). In the current study, the uninjured observations were randomly undersampled by 50% and the injured observations were synthetically oversampled by 100%. Predictive models were built using SMOTE and no SMOTE to compare performance.

Cross-validation and parameter tuning. Supervised learning algorithms have a number of parameters (see Table, Supplemental Digital Content 1, A brief description of each machine learning algorithm and all relevant parameters, http://links.lww.com/MSS/B163) which can be tuned to determine how an algorithm interacts with the training data (25). The aim of supervised learning is to select the parameters that optimise an algorithm's ability to perform on the testing data (25). The most common method of finding the optimal parameter combination is by performing a grid search that considers all parameter combinations (25). One potential issue with this method however, is overfitting, which occurs when the parameters selected are too closely fit to the training data from which they were derived (26). This reduces generalisability and the ability of an algorithm to perform on the testing data (26). A solution to overfitting is cross-validation (25). Tenfold cross-validation splits the training data into 10 equal subsets. One of the subsets is withheld, and the remaining nine subsets are used to perform a grid search for the optimal parameter combination. The selected parameters are then validated using the withheld subset. This process is repeated tenfold, with each subset being withheld once as the validation subset. The parameter combination that performs the best across each fold is then selected for the final model. Tenfold cross validation was used in the current study to select the optimal parameter combination for each algorithm.

Performance outcomes. Accuracy can be a poor indicator of performance when attempting to predict injury incidence (24). Cohen's kappa coefficient is an alternative to accuracy that accounts for the base rate of expected accuracy due to random chance. However, both accuracy and kappa are calculated based on the number of correct and incorrect binary classifications and do not take into account the estimated probability of an observation belonging to one class or the other. For example, if athlete A has a 49% probability of HSI and athlete B has 1% probability of HSI, both these athletes are predicted to remain uninjured. Accuracy and kappa do account for the fact that athlete A was 48% more likely to sustain an HSI than athlete B. Area under the curve (AUC) of a receiver operator characteristic curve measures the probability that a positive case will be ranked higher than a negative case. An AUC of 0.5 indicates prediction no better than random chance, with a value of 1.0 indicating perfect prediction (13). In the current study, predicted injury probabilities were used to construct a receiver operator characteristic curve, and AUC was used to measure the likelihood that the observations with a higher probability of injury were actually the injured observations. As AUC is calculated using predicted injury probabilities and not the number of correct and incorrect binary classifications, it can be a more sensitive measure of performance than accuracy or kappa (27). Accordingly, the AUC of each algorithm was calculated and was used to compare predictive performance as well as to select the optimal parameter combination for each model.

RESULTS

Cohort and Prospective Injury Details

2013. One-hundred eighty-six athletes (age, 23.2 ± 3.6 yr; stature, 188.0 ± 7.1 cm; mass, 87.6 ± 7.5 kg) participated in the 2013 study period. Of these, 27 sustained a prospective HSI (age, 23.8 ± 3.6 yr; stature, 185.3 ± 6.3 cm; mass, 84.4 ± 5.6 kg) and 159 did not (age, 23.1 ± 3.6 yr; stature, 188.4 ± 7.2 cm; mass, 88.1 ± 7.7 kg). High-speed running was the most common mechanism of injury (59%), and biceps femoris long head was the most commonly injured muscle (78%). Five injuries occurred during the preseason period, and 22 occurred during the in-season period. The average eccentric hamstring strength and between-limb imbalance of the injured athletes was 260 ± 79 N and 45 ± 46 N respectively. For the uninjured athletes, the average eccentric hamstring strength and between-limb imbalance was 301 ± 84 N and 49 ± 56 N, respectively.

2015. During the 2015 study period, 176 athletes participated (age, 25.0 ± 3.4 yr; stature, 187.6 ± 7.5 cm; mass, 87.0 ± 8.6 kg). Twenty-six athletes sustained a HSI (age, 25.2 ± 3.4 yr; stature, 187.8 ± 7.3 cm; mass, 87.1 ± 8.3 kg) and 150 did not (age, 25.1 ± 3.5 yr; stature, 187.7 ± 7.8 cm; mass, 87.0 ± 8.9 kg). The most common mechanism of HSI was high-speed running (50%), and the most commonly injured muscle was biceps femoris (92%). Nine injuries occurred during the preseason period, and 17 occurred during the in-season period. The average eccentric hamstring strength and between-limb imbalance of the injured athletes was 341 ± 80 N and 30 ± 21 N, respectively. For the uninjured athletes, the average eccentric hamstring strength and between-limb imbalance was 341 ± 78 N and 34 ± 30 N, respectively.

Within-Year Predictive Performance

2013. The median AUC for all 2013 within-year predictive models was 0.58. There were no large differences between the predictive performance of models built using all variables as predictors (median AUC of 0.57) and models built using only the three variables as predictors (median AUC of 0.59). Using SMOTE and no SMOTE input data did not improve predictive performance, with both sets of models resulting in a median AUC of 0.58 and 0.59, respectively. For all 2013 within-year predictive models, AUC ranged from 0.26 to 0.91. Naive Bayes was the best-performing algorithm, with a median AUC of 0.60. The performance of each individual algorithm for 2013 within-year predictive models

built using three variables, and all variables can be found in Figures 2A and B, respectively.

2015. For all 2015 within-year predictive models, the median AUC was 0.57. Similar to 2013, there was no difference in the predictive performance of models built using all variables as predictors and only the three variables as predictors, with both sets of models resulting in a median AUC of 0.57. The case was the same for models using SMOTE and no SMOTE input data (median AUC of 0.57). The range in AUC for the 2015 models was slightly larger than the 2013 models (0.24 to 0.92). The equal best-performing algorithms were random forest and support vector machine, with a median AUC of 0.58. The performance of each individual algorithm for 2015 within-year predictive models built using three variables and all variables can be found in Figures 3A and B, respectively.

Between-Year Predictive Performance

The performance of the between-year predictive models was poorer than the within-year predictive models, with a median AUC of 0.52. The median AUC was 0.52 and 0.53 for models built using all variables and three variables, respectively. Similar to the within-year predictive models,

using SMOTE input data did not significantly improve performance, with these models resulting a median AUC of 0.53 compared with a median AUC of 0.52 for models built using no SMOTE input data. There was little variation in AUC for the independent algorithms as each iteration was performed using the same training and testing data. Across all the between-year predictive models, AUC ranged from 0.37 to 0.73. Naïve Bayes was the best-performing algorithm, with a median AUC of 0.54. The performance of each individual algorithm for between-year predictive models built using three variables and all variables can be found in Table 1.

DISCUSSION

The aim of the current study was to investigate whether the application of various supervised learning techniques, using risk factor and demographic data collected over two seasons, could be used to predict the occurrence of HSI. The main finding of this study was that eccentric hamstring strength, age, and previous HSI data cannot be used to identify athletes at an increased risk of HSI with any consistency. Although some iterations of the within-year predictive models achieved near perfect performance (maximum AUC
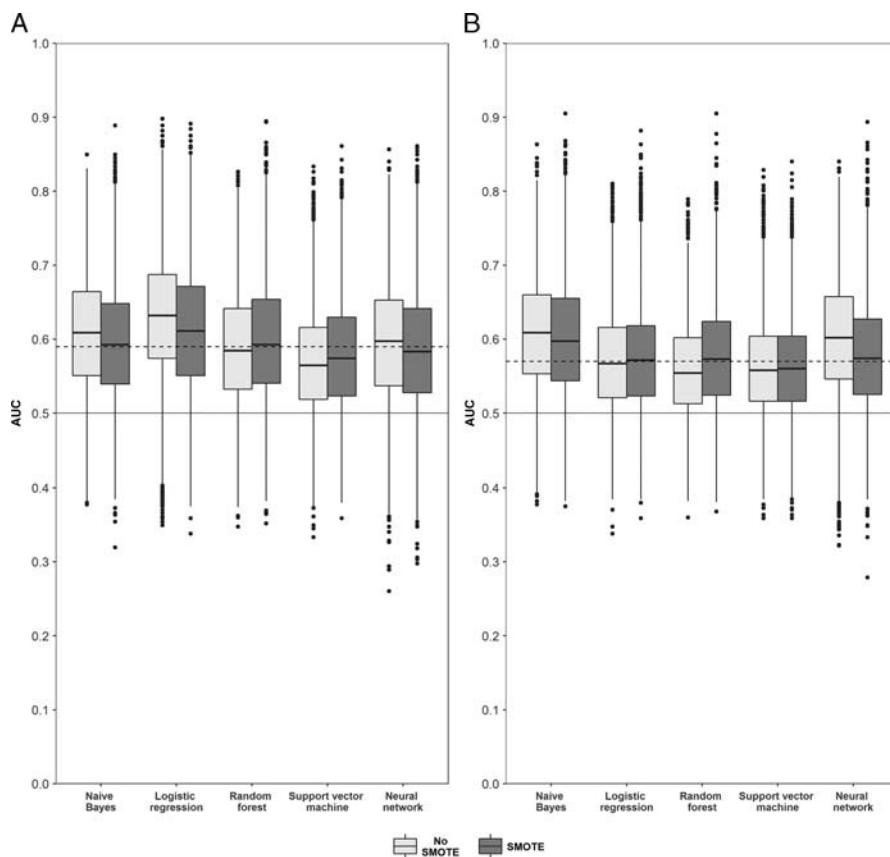

FIGURE 2—The individual performance of 2013 within-year HSI prediction models built using (A) eccentric hamstring strength, age and previous HSI as input data, and (B) eccentric hamstring strength, age, previous HSI, between-limb imbalance, previous anterior cruciate ligament injury, stature, mass and primary playing position as input data. The *solid horizontal line* indicates the expected performance due to random chance and the *dotted horizontal line* represents the median AUC of all models.
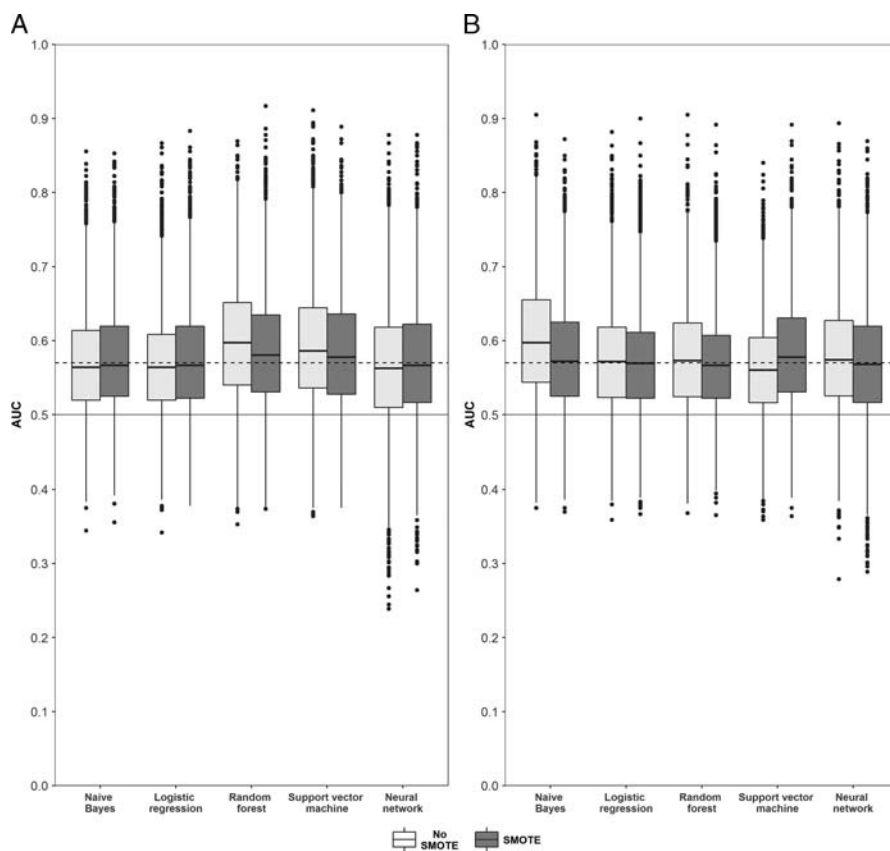
FIGURE 3—The individual performance of 2015 within-year HSI prediction models built using **(A)** eccentric hamstring strength, age and previous HSI as input data, and **(B)** eccentric hamstring strength, age, previous HSI, between-limb imbalance, previous anterior cruciate ligament injury, stature, mass, and primary playing position as input data. The *solid horizontal line* indicates the expected performance due to random chance and the *dotted horizontal line* represents the median AUC of all models.

of 0.92), others performed worse than random chance (minimum AUC of 0.24).

The large discrepancy in AUC is indicative of the fragility of the data. Although each iteration was performed using the same data, small changes in the randomly sampled training and testing data vastly influenced the performance of the within-year predictive models. This not only demonstrates the sensitivity of AUC to sampling but also suggests that larger data sets are needed when investigating the predictive ability of injury risk factors. Collecting more data will likely result in an increase in the number of observed injuries and consequently a more robust data set. This will in turn improve the ability of supervised learning techniques to identify patterns with more consistency, should such patterns exist. In addition to within-year predictive performance, this study also investigated the performance of between-year predictive models. As the training data (2013 data) and testing data (2015 data) were the same for every iteration, the resulting AUC ranges were much smaller than within-year predictive models, with any variability in AUC only caused by cross-validation and SMOTE. Despite less discrepancy, the performance of between-year predictive models was poor, with a median AUC of 0.52.

The etiology of HSI is multifactorial, and injuries typically occur as a result of the interactions between numerous variables (28,29). The poor predictive performance displayed by between-year predictive models may be due to differing contributing factors between the 2013 and 2015 injuries, specifically the role of eccentric hamstring strength. In 2013, the injured athletes were, on average, 42 N weaker than the uninjured athletes. In 2015, however, there was no difference in the average eccentric hamstring strength between the injured and uninjured athletes. Overall, the 2015 cohort was stronger than the 2013 cohort. Despite this, the percentage of players that sustained an HSI was identical in both seasons (15%), which suggests that eccentric hamstring strength was less of an influencing factor in the HSI that occurred in 2015, as opposed to 2013. Previous work, which investigated eccentric hamstring strength and HSI risk, found that weaker rugby union players were no more likely to sustain an HSI than stronger players (20). The authors of this study (20) suggest that the conflicting results with previously published work (10) may be due to the rugby union players investigated (20) being considerably stronger than the previously investigated Australian footballers (10). It is also suggested that athletes with high levels of eccentric hamstring strength may not see any additional protective benefit from further increases in strength (10,20). This hypothesis aligns with the current findings that eccentric hamstring strength may not play as large a role in HSI risk

TABLE 1. Descriptive statistics summarising the individual performance of between-year HSI prediction models built using 2013 data for training and 2015 data for testing.

		AUC			
		Group 1[a]		Group 2[b]	
Algorithm		No SMOTE	SMOTE	No SMOTE	SMOTE
Naïve Bayes	Minimum	0.49	0.41	0.47	0.42
	Lower quartile	0.54	0.53	0.47	0.52
	Median	0.54	0.55	0.47	0.55
	Upper quartile	0.54	0.56	0.47	0.57
	Maximum	0.54	0.65	0.57	0.65
Logistic regression	Minimum	0.54	0.45	0.53	0.43
	Lower quartile	0.54	0.48	0.53	0.50
	Median	0.54	0.54	0.53	0.52
	Upper quartile	0.54	0.54	0.53	0.54
	Maximum	0.54	0.58	0.53	0.62
Random forest	Minimum	0.48	0.42	0.45	0.41
	Lower quartile	0.51	0.50	0.52	0.51
	Median	0.51	0.52	0.53	0.53
	Upper quartile	0.52	0.54	0.54	0.55
	Maximum	0.54	0.65	0.57	0.65
Support vector machine	Minimum	0.43	0.41	0.45	0.41
	Lower quartile	0.49	0.51	0.48	0.50
	Median	0.51	0.53	0.52	0.52
	Upper quartile	0.52	0.55	0.52	0.55
	Maximum	0.61	0.68	0.54	0.66
Neural network	Minimum	0.44	0.37	0.45	0.40
	Lower quartile	0.50	0.50	0.50	0.50
	Median	0.51	0.52	0.50	0.52
	Upper quartile	0.54	0.54	0.51	0.54
	Maximum	0.65	0.65	0.57	0.73

Performance is measured as AUC.
[a]Group 1; models built using eccentric hamstring strength, age and previous HSI as input data.
[b]Group 2; models built using eccentric hamstring strength, age, previous HSI, between-limb imbalance, previous anterior cruciate ligament injury, stature, mass and primary playing position as input data.
SMOTE, synthetic minority oversampling techniques.

in a stronger cohort (2015) compared with a weaker cohort (2013).

The current study investigated whether the inclusion of data that have not been directly linked to HSI risk in Australian footballers, in addition to previously reported HSI risk factors, improved predictive performance. The inclusion of between-limb imbalance, previous ACL injury, stature, mass and playing position, in addition to eccentric hamstring strength, age, and previous HSI, did not improve predictive performance (Figs. 2 and 3). In some cases, these data may have confounded predictive performance, with the majority of the 2013 within-year predictive models performing better when built using only previously reported HSI risk factors (Fig. 2). The results of the current study are comparable to the findings of prior work, which attempted to predict all injuries using workload data from a single AFL team (30). This study observed a mean AUC of 0.65 when predicting all noncontact injuries. Hamstring-related injuries were predicted with a mean AUC of 0.72; however, these injuries were not specifically HSI (30). Earlier work has observed an association between high-speed running distances and HSI risk (8,11); however, it was concluded that these measures examined in isolation are poor predictors of HSI and should be examined in concert with other variables (11). The results of previous work (8,11,30) suggest that the inclusion of workload data, in addition to other HSI risk factors, may improve the ability to predict HSI occurrence; however, this is yet to be examined.

Imbalanced classes have previously been highlighted as a limitation of injury prediction research (30). This study compared the performance of predictive models built using SMOTE and no SMOTE. Using SMOTE to undersample the uninjured observations and oversample the injured observation did not improve predictive performance for within-year predictive models (Figs. 2 and 3) or between-year predictive models (Table 1). These results are in contrast to prior work, which observed increases in predictive performance when SMOTE were used to build support vector machine models (30). It is possible that SMOTE is only beneficial when used to synthetically reproduce predictor variables from a complex data set, such as workload derived from global position systems data. In the current study, the uninjured observations were randomly undersampled by 50%, and the injured observations were synthetically oversampled by 100%. However, these values were chosen arbitrarily, and it is unknown whether a different percentage of undersampling and oversampling would impact predictive performance.

The current study has a number of limitations that may have influenced predictive performance. The data used as predictor variables in the current study were only collected at the beginning of preseason training for each study period. It is unknown whether more frequent measures of variables, such as eccentric hamstring strength, would have improved predictive performance. It is also unknown whether the magnitude of change in eccentric hamstring strength across a season may be a more sensitive measure than absolute strength. Although eccentric hamstring strength, age, and previous HSI are purported as HSI risk factors (9,10,12,29,31), there are a number of other factors not included in this study which have been linked to the risk of HSI.

The inclusion of further HSI risk factors, such as workload (8,11) and biceps femoris fascicle length (21), may improve the ability to predict HSI.

Furthermore, there is a lack of literature regarding the application of supervised learning techniques in injury prediction research. The proportions of data used as training and testing data, as well as the proportions of data undersampled and oversampled, were chosen arbitrarily, and it is unknown whether different proportions would have influenced predictive performance. There are a number of different supervised learning techniques, and although some of the techniques applied in the current study have been applied in previous work (30), it is unknown how different techniques would have impacted predictive performance. Lastly, the current findings relate to HSI that fulfilled the criteria of acute pain in the posterior thigh that resulted in disruption of the hamstring fibers, as confirmed by MRI (10). Cases that were MRI negative but clinically positive were not included in this study, and it is difficult to determine how the inclusion of these injuries would have impacted the findings. In addition to this, other injury types are certain to occur, and these should be accounted for by including them as competing risks in the modeling approach (32).

In conclusion, eccentric hamstring strength, age and previous HSI data cannot be used to identify Australian footballers at an increased risk of HSI with any consistency. Despite suggestions that the ability to predict injury with prospectively collected data is limited, the application of supervised learning techniques in elite sport still warrants further investigation. It is unclear whether injury probabilities calculated by predictive models can be used to identify at risk athletes and guide risk mitigation strategies to ultimately reduce injury incidence. With the previously discussed considerations in mind, future research is needed to determine what, if any, improvements can be made in injury prediction performance.

REFERENCES

1. Orchard JW, Seward H, Orchard JJ. Results of 2 decades of injury surveillance and public release of data in the Australian Football League. Am J Sports Med. 2013;41(4):734–41.
2. Fyfe JJ, Opar DA, Williams MD, Shield AJ. The role of neuromuscular inhibition in hamstring strain injury recurrence. J Electromyogr Kinesiol. 2013;23(3):523–30.
3. Verrall GM, Kalairajah Y, Slavotinek JP, Spriggins AJ. Assessment of player performance following return to sport after hamstring muscle strain injury. J Sci Med Sport. 2006;9(1–2):87–90.
4. Opar DA, Williams MD, Timmins RG, Dear NM, Shield AJ. Knee flexor strength and bicep femoris electromyographical activity is lower in previously strained hamstrings. J Electromyogr Kinesiol. 2013;23(3):696–703.
5. Hickey J, Shield AJ, Williams MD, Opar DA. The financial cost of hamstring strain injuries in the Australian Football League. Br J Sports Med. 2014;48(8):729–30.
6. Bennell K, Tully E, Harvey N. Does the toe-touch test predict hamstring injury in Australian Rules footballers? Aust J Physiother. 1999;45(2):103–9.
7. Brockett CL, Morgan DL, Proske U. Predicting hamstring strain injury in elite athletes. Med Sci Sports Exerc. 2004;36(3):379–87.
8. Duhig S, Shield AJ, Opar D, Gabbett TJ, Ferguson C, Williams M. Effect of high-speed running on hamstring strain injury risk. Br J Sports Med. 2016;50:1536–40.
9. Gabbe BJ, Bennell KL, Finch CF, Wajswelner H, Orchard JW. Predictors of hamstring injury at the elite level of Australian football. Scand J Med Sci Sports. 2006;16(1):7–13.
10. Opar DA, Williams MD, Timmins RG, Hickey J, Duhig SJ, Shield AJ. Eccentric hamstring strength and hamstring injury risk in Australian footballers. Med Sci Sports Exerc. 2015;47(4):857–65.
11. Ruddy JD, Pollard CW, Timmins RG, Williams MD, Shield AJ, Opar DA. Running exposure is associated with the risk of hamstring strain injury in elite Australian footballers. Br J Sports Med. 2016; [Epub ahead of print]. doi:10.1136/bjsports-2016-096777.
12. Verrall GM, Slavotinek JP, Barnes PG, Fon GT, Spriggins AJ. Clinical risk factors for hamstring muscle strain injury: a prospective study with correlation of injury by magnetic resonance imaging. Br J Sports Med. 2001;35(6):435–9.
13. Bahr R. Why screening tests to predict injury do not work—and probably never will…: a critical review. Br J Sports Med. 2016;50:776–80.
14. Bittencourt NF, Meeuwisse WH, Mendonca 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 Sports Med. 2016;50:1309–14.
15. Mendiguchia J, Alentorn-Geli E, Brughelli M. Hamstring strain injuries: are we heading in the right direction? Br J Sports Med. 2012;46(2):81–5.
16. Quatman CE, Quatman CC, Hewett TE. Prediction and prevention of musculoskeletal injury: a paradigm shift in methodology. Br J Sports Med. 2009;43(14):1100–7.
17. Han J, Kamber M. Chapter 6: Classification and Prediction. Data Mining: Concepts and Techniques. 2nd ed. San Francisco: Elsevier; 2006. pp. 285–378.
18. Bahr R, Holme I. Risk factors for sports injuries—a methodological approach. Br J Sports Med. 2003;37(5):384–92.
19. van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries. A review of concepts. Sports Med. 1992;14(2):82–99.
20. Bourne MN, Opar DA, Williams MD, Shield AJ. Eccentric knee flexor strength and risk of hamstring injuries in Rugby Union: a prospective study. Am J Sports Med. 2015;43(11):2663–70.
21. Timmins RG, Bourne MN, Shield AJ, Williams MD, Lorenzen C, Opar DA. Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. Br J Sports Med. 2016;50:1524–35.

22. R Core Team. *R: A Language and Envrionement for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013.

23. Japkowicz N, Stephen S. The class imbalance problem: a systematic study. *Intell Data Anal*. 2002;6(5):429–49.

24. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321–57.

25. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res*. 2012;13:281–305.

26. Kotsiantis SB. Supervised machine learning: a review of classification techniques. *Informatica*. 2007;31:249–68.

27. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit*. 1997;30(7): 1145–59.

28. Croisier JL. Factors associated with recurrent hamstring injuries. *Sports Med*. 2004;34(10):681–95.

29. Opar DA, Williams MD, Shield AJ. Hamstring strain injuries: factors that lead to injury and re-injury. *Sports Med*. 2012;43(3):209–26.

30. Carey DL, Ong K-L, Whiteley R, Crossley KM, Crow J, Morris ME. Predictive modeling of training loads and injury in Australian football. arXiv. 2017; arXiv:1706.04336.

31. Freckleton G, Pizzari T. Risk factors for hamstring muscle strain injury in sport: a systematic review and meta-analysis. *Br J Sports Med*. 2013;47(6):351–8.

32. Nielsen RO, Malisoux L, Moller M, Theisen D, Parner ET. Shedding light on the etiology of sports injuries: a look behind the scenes of time-to-event analyses. *J Orthop Sports Phys Ther*. 2016; 46(4):300–11.

CLINICAL SCIENCES