Richter, C., Petushek, E., Grindem, H., Franklyn-Miller, A.,Bahr, R., Krosshaug, T. (2021). Cross-validation of a machine learning algorithm that determines anterior cruciate ligament rehabilitation status and evaluation of its ability to predict future injury. *Sports Biomechanics.* http://dx.doi.org/10.1080/14763141.2021.1947358

Cross-validation of a machine learning algorithm that determines anterior cruciate ligament rehabilitation status and evaluation of its ability to predict future injury.

Chris Richter [1, 2], Erich Petushek[3], Hege Grindem[4,5] Andrew Franklyn-Miller[1,6], Roald Bahr[4,7] & Tron Krosshaug[4]

1 Sports Medicine Department, Sports Surgery Clinic, Santry Demesne, Ireland

2 Department of Life Sciences, Roehampton University, UK

3 Department of Cognitive and Learning Sciences, Michigan Technological University, USA

4 Oslo Sport Trauma Research Center, Department of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway

5 Stockholm Sports Trauma Research Center, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

6 Centre for Health, Exercise and Sports Medicine, University of Melbourne, Australia

7 Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar

Corresponding author: Dr Chris Richter
Head of Data and Innovation
SSC Sports Medicine
Unit C10 Gulliver's Retail Park, Northwood, Santry, Dublin 9
Tel: +353 83 4137 223
Fax: + 353.1.5262046

Keywords: Classification algorithms, Injury prediction, vertical drop jump

Competing interests: **No, there are no competing interests for any author.**

Contributions: **All authors were part of the content creation, the writing and editing process of the manuscript.**

Funding info: **No funding was received by the authors for the work.**

Ethical approval information **Data capture and analysis was granted ethical approval of the corresponding institution as stated in previous publications.**

Data sharing statement **The data that support the findings of this study are available from the corresponding author, CR and TK, upon reasonable request.**

Cross-validation of a machine learning algorithm that determines anterior cruciate ligament rehabilitation status and evaluation of its ability to predict future injury.

Abstract

Classification algorithms determine the similarity of an observation to defined classes, e.g., injured or healthy athletes, and can highlight treatment targets or assess progress of a treatment. The primary aim was to cross-validate a previously developed classification algorithm to identify rehabilitation status [Richter et al. (2019)] using a different sample [Krosshaug et al. (2016)], while a secondary aim was to examine its ability to predict future ACL injuries. The examined outcome measure was "healthy-limb" class membership probability, which was compared between a cohort of athletes without previous or future (No Injury), previous (PACL) and future ACL injury (FACL). The No Injury group had significantly higher probabilities than the PACL ($p<0.001$; medium effect) and FACL group ($p\leq0.045$; small effect). The ability to predict group membership was poor for the PACL (area under curve [AUC]; 0.61<AUC<0.62) and FACL group (0.57<AUC<0.59). The ACL injury incidence proportion was highest in athletes with probabilities below 0.20 (9.4%; +2.7% to baseline). Athletes with probabilities between 0.20 and 0.80 had an incidence proportion of 6.5% (-0.2%), while athletes with probabilities above 0.80 had an incidence proportion of 4.1% (-2.6%). While findings that a low probability might represent an increase in injury risk on a group level, it is not sensitive enough for injury screening to predict a future injury on the individual level.

Introduction

Classification algorithms determine the similarity of an observation to defined classes, e.g., injured or healthy athletes. These algorithms have a big impact on medical practice (Hannun et al., 2019) and help health care practitioners by highlighting treatment targets or assessing progress with treatment (Barton, Hawken, Scott, & Schwartz, 2012). In medical imaging, classification algorithms have outperformed practitioners in detecting disease (Hannun et al., 2019), and in gait analysis, algorithms have been reported to objectively describe the deviation of an individual's movement from normality (Barton et al., 2012). Another example is a recent study on male field sport athletes, by Richter et al. (2019) that presented a classification algorithm that was trained to distinguish athletes with anterior cruciate ligament (ACL) reconstruction from uninjured athletes using biomechanical features extracted from the vertical drop jump data (healthy males from a variety of multidirectional field sports and males who had an ACL reconstruction approximately 9 months before testing). The classification algorithm was judged to be able to estimate rehabilitation status because it was able to accurately differentiate between healthy, operated, and limbs contralateral of an operated limb (area under the curve, AUC: 0.94, measure was generated post-publication by the study authors). Such feedback could be beneficial to coaches, physiotherapists and other medical professionals in order to determine the player's rehabilitation state, i.e., whether or not he/she is ready to return to sport. Furthermore, such a classification algorithm might even be able to objectively estimate ACL injury risk because athletes with previous ACL injury have an up to 15 times higher risk for a (second) ACL injury (Paterno, Rauh, Schmitt, Ford, & Hewett, 2014; Wiggins et al., 2016), which may be mediated, at least partly, by altered movement patterns. However, it is unclear how the classification algorithm developed by Richter et al. (2019) would perform in other cohorts and if athletes with a low similarity to the "healthy limb" class have a greater risk for sustaining a future ACL.

Therefore, the primary aim of this investigation was to cross-validate the classification algorithm by Richter et al. (2019) in its ability to assess ACL rehabilitation status, using a cohort of female handball and football players. A secondary aim was to assess the classification algorithm's ability to predict future risk of ACL injury. We hypothesized a healthy individual will have higher probabilities to belong to the control class than previously injured individuals and individuals that will suffer an injury in the future.

Methods

This study used kinematic and kinetic measures of a double leg drop vertical jump from two previously published studies - Richter et al. (2019) and Krosshaug et al. (2016). Analyses were performed in three steps: 1) Building the classification algorithm, 2) Classification (using the classification algorithm), and 3) Statistical analyses (testing the classification algorithm). When building the classification algorithm, double leg drop vertical drop jump data without an arm-swing from 218 males from the Richter et al. (2019) study were used, while double leg drop vertical drop jump data with an arm swing from 822 females from the Krosshaug et al. (2016) study were used for the cross-validation of the classification algorithm and the future injury prediction analysis.

Building the Classification Algorithm

The used classification algorithm was built by following three steps: feature extraction (data driven extraction), machine learning technique selection (identification of the technique that is most suitable for the classification task using all extracted features) and identification of the most meaningful features (to generate an interpretable classification model). Richter et al. (2019) describe in detail how prediction features were extracted, how kinematic measures were obtained and why a pattern recognition network was used. To generate the classification algorithm, a pattern recognition network was trained to classify an input matrix (training matrix; 3 variables [features] x 468 observations) into 3 target classes (healthy, ACL operated and contralateral limb) utilizing only data, features and processes described in Richter et al. (2019). The class “healthy” was represented by the left and right limb of athletes with no ACL injury. The class “ACL operated” was represented by the limb with a reconstructed ACL. The class “contralateral” was represented by the contralateral limb of athletes with reconstructed ACL. In contrast to the Richter et al. (2019) study, the sample size of the healthy limb class was artificially increased by 32 (increasing the sample size of observations in the input matrix to 500 observations) to balance the number of observations across the classes and thereby eliminate effects of unbalanced sample sizes during the fitting process. This was done by using synthetic minority over-sampling technique. During the synthetic minority over-sampling, only observations that were never misclassified across 100 simulations were used to increase the sample of the healthy class (this information was taken from unpublished findings of Richter et al., 2019) to purify the class characteristics that separated the healthy limb class from the ACL operated and contralateral limb class. Consequently, the classification algorithm was generated utilizing 156 observations for the healthy class (62 male athletes [124 limbs] and 32 synthetically engineered limbs) and 156 observations for the ACL operated and contralateral classes (156 athletes who underwent primary ACL reconstruction approximately 9 months before the biomechanical assessment). Features used in the training matrix were: the average magnitude of the vertical ground reaction force asymmetry from 80 to 86 % of the movement cycle (feature: vGRF asymmetry), the average magnitude of the resultant centre of mass velocity from 80 to 84 % of the movement cycle (feature: CoM velocity) and the average knee flexion angular velocity asymmetry from 74 to 81 % of the movement cycle (feature: knee velocity asymmetry). The asymmetry features (vGRF asymmetry and knee velocity asymmetry) were computed as reported in equation 1.

$$a\text{symmetry feature} = \begin{cases} feature_{left} - feature_{right}; & if\ side = left \\ feature_{right} - feature_{left}; & if\ otherwise \end{cases} \qquad \text{equation 1}$$

Before model fitting, all 3 features were transformed into z-scores by dividing the centred scores (score–pooled mean of whole matrix) by its pooled standard deviation.

Applying the Classification Algorithm to a New Dataset

The classification algorithm was applied to data previously published in Krosshaug et al. (2016), which describes in detail how data was collected and processed. After the classification algorithm was built using the data of Richter et al. (2019), it was used to classify the data from the Krosshaug et al. (2016) study. These data consisted of vertical drop jump with arm-swing recorded from 822 female athletes. Athletes were assigned to the following cohorts: no previous or future ACL injury (No Injury; n = 693), previous ACL injury (PACL; n = 74) and future ACL injury (FACL; n = 55, which can be sub-grouped into 1st ACL injury [n = 43] and 2nd ACL injury [n = 12]). A subject was assigned 2nd ACL injury class in the case of either an ACL graft rupture or a contralateral ACL injury. Before extracting the required features from the data, a selection of kinematic and kinetic measures was screened for errors using a custom-written graphical user interface to ensure the same processing flow as the dataset used to build the classification algorithm. Subsequently, a 3x4932 input matrix (testing matrix; 4932 observations = 3 trials x 2 body sides x 822 subjects) was built. Only one limb was randomly selected for each athlete in the no previous or future ACL injury class. This testing matrix was then normalized using the mean and standard deviation observed during the model generation and fed into the classification algorithm, which computed class probability scores for each observation within the testing matrix. Only the probability of belonging to the healthy limb class (healthy score) was used for further analysis.

Cross-Validation / Score Analysis

To examine if there were differences in the magnitude of the healthy score between the defined cohorts, a one-sided Mann-Whitney-Wilcoxon test, treating the scores as non-parametric and unpaired, was performed on the healthy scores. As each athlete performed multiple vertical drop jumps and the selection of a specific trial could impact the results, the score analysis was performed using 1) the average healthy score across trials, 2) the maximal trial (highest jump) or 3) a random healthy score from one of the three trials. For the No Injury cohort, we defined the average healthy score as the average of all performed trials from a random side (left or right). For the other cohorts, only the side that suffered or had previously suffered an ACL injury was used. We used the same approach to analyse the maximal trial, using only the vertical drop jump trial with the highest jump height. For the random healthy score analysis, 100 random analyses (shuffle splits) were performed. For each shuffle split, a random trial was chosen from a random side (for the No Injury group) or of the side that suffered the ACL injury (for other groups). To judge the effect size of the difference, we computed Cohens D from the Mann-Whitney-Wilcoxon U_2 score (Borenstein, Hedges, Higgins, & Rothstein, 2009; Kerby, 2014) and classified the effect as large ($d > 0.80$), medium ($0.80 > d \geq 0.50$), small ($0.50 > d \geq 0.20$) and negligible ($d < .20$; Cohen, 1988).

To examine the predictive ability of the classification model, a receiver-operating curve (ROC) was computed. The area under the curve (AUC) was used to classify the combined sensitivity and specificity as outstanding ($0.90 \geq AUC < 1.00$), excellent ($0.80 \geq AUC < 0.90$), acceptable ($0.70 \geq AUC < 0.80$), poor ($0.50 \geq AUC < 0.80$) and no discrimination ($AUC < 0.50$; Hosmer et al., 2013). The receiver-operating curve (ROC) was computed for comparisons between the defined cohorts: No Injury and PACL, and No Injury and FACL [1st and 2nd ACL injury]. Average positive predictive value (precision) was computed and used as a measure of the ability of the healthy scores to classify PACL from No Injury as well as FACL from No Injury. Sensitivity and specificity measures were reported at the optimal operating point of the ROC curve, which was defined as the point where the Youden's index was maximized (Youden, 1950).

The alpha level of this study was set to 0.05. No adjustments for multiple comparisons were made (Rothman, 1990). No classification of sensitivity, specificity or precision was given as corresponding thresholds are dependent on the “use” case, which includes the individual appraisal of costs/benefits. All analyses were performed with MATLAB 2018a.

Results

The defined subgroups of the Richter et al. (2019) and Krosshaug et al. (2016) data demonstrated some differences in respect to the extracted features: vGRF asymmetry, knee velocity asymmetry and CoM velocity (table 1). Compared to the testing matrix, the training matrix had larger vGRF asymmetry, larger standard deviations in the knee velocity asymmetry and higher CoM velocity.

Table 1. Mean (standard deviation) values of the three classification features (vertical ground reaction force asymmetry [vGRF asymmetry], resultant centre of mass velocity [CoM velocity] and knee flexion angular velocity symmetry [knee velocity asymmetry]) in athletes with no previous or future ACL injury (No Injury), previous ACL injury (PACL), future ACL injury (FACL).

The No Injury group had, on a group-level, significantly higher healthy scores than the PACL group (average analysis: $p < 0.001$; $d = 0.53$, maximal analysis: $p < 0.001$: $d = 0.45$; 100 of 100 shuffle splits [$d = 0.44$ to 0.45]; figure 1; table 2). The ability to predict group membership was poor (average analysis: AUC = 0.62; maximal analysis: AUC = 0.61), with the average precision being over baseline (average analysis: +3%; maximal analysis: +4%). The sensitivity was 0.59 in the average and 0.61 in the maximal analysis, while specificity was 0.58 and 0.61, respectively.

Table 2. Median (interquartile range [IQR]) of healthy score magnitudes and findings when comparing the no previous or future ACL injury (No Injury) and previous ACL injury (PACL) group. The measure: “n split” refers to the number of differences observed during the shuffle split analysis.

Figure 1. On the left Illustration for the maximal healthy scores in athletes with no previous or future ACL injury (No Injury), previous ACL injury (PACL), future first ACL injury (1st ACL injury) and future second ACL injury (2nd ACL injury). On the right illustration of ACL injury / arbitrary proportion across selected ranges of the healthy score, computed using bootstrap samplings. The ’*’ symbol indicates a significant difference ($p < .001$) of the observed incidence proportion to the baseline incidence proportion based on the average p value resulting of 100 sign-tests utilizing bootstrap samplings (bootstraps = sample size of the athletes included within the range).

No significant differences were found on a group level between the No Injury and 1st ACL cohort for the average scores (p = 0.181; *d* = 0.16; figure 1), while the No Injury cohort had significantly higher healthy scores in the maximal analysis (p = 0.045; *d* = 0.30: figure 1) and some shuffle splits (3 of 100; d = 0.18 to 0.19; table 3). When comparing the No Injury group and FACL [1st and 2nd ACL injury], healthy scores were significantly higher in the No Injury group (average analysis: p = 0.045; *d* = 0.28: maximal analysis: p = 0.011; *d* = 0.36; 78 of 100 splits; d = 0.28 to 0.29; figure 1; table 3).

Table 3. Median (interquartile range [IQR]) and findings when comparing the no previous or future ACL injury (No Injury) and future first ACL injury (1st ACL) group as well as No Injury and future ACL injury (FACL). The measure: "n split" refers to the number of differences observed during the shuffle split analysis.

The future injury rate (incidence proportion) across the entire sample was 6.7%. When examining arbitrary proportion across selected ranges of the healthy score, athletes with healthy scores below 0.33 had a future ACL injury incidence proportion of 7.6% (n = 408; 5th to 95th percentile; 5.9 to 12.1%). Athletes with healthy scores between 0.33 and 0.66 had an ACL injury incidence proportion of 6.1% (n = 161; 3.0 to 12.1%), while athletes with healthy scores above 0.66 demonstrated an ACL injury incidence proportion of 5.5% (n = 253; 3.5 to 9.5%; table 3; figure 1).

The ability to predict future ACL injuries was poor in the average and maximal analysis (AUC = 0.57 and AUC = 0.59; figure 2), with average precisions slightly over baseline (+2% and +3%). The sensitivity was 0.58 in the average and 0.56 in the maximal analysis, while specificity was 0.53 and 0.53, respectively.

Figure 2. Illustrated on the left side are precision and recall curves (and the resulting confusion matrix) for the classification tasks and on the right side are the receiver operation curve curves for the maximal scores. The red line displays the sensitivity and specificity for the classification between the groups: no previous or future ACL injury (No Injury) and previous ACL injury (PACL), while the black line displays: No Injury and future ACL injury (FACL).

Discussion and Implication

Main Findings

This study cross-validated a classification algorithm previously developed to identify athletes in male rugby, football and Gaelic games with previous ACL reconstruction using a countermovement jump without arm-swing, in a different sample of female elite handball and football players that performed a countermovement jump with arm-swing. Additionally, we examined if the outcome measures of classification algorithm (class probability scores) could identify players with increased risk for a future ACL injury. The results demonstrated that athletes in the cross-validation cohort with either a previous (PACL) or a future injury (FACL) had lower healthy scores than athletes without previous or future injury (No Injury). However, the classification accuracy was poor in both cases. The poor classification performance was caused to a large portion by a misclassification of samples from the No Injury group (classified into ACL class), a high number of cases [50 %] demonstrating a healthy score > .50 (binominal distribution peak at 0.1 and 0.8), while misclassifications in the PACL and FACL cohort (classified into Healthy class) were fewer. In the PACL and FACL cohort, about 30 % of the samples demonstrated a score > .50 (unimodal skewed distribution peak at > 0.1). This is similar to findings of Richter et al. (2019) where the No Injury cohort was also more often misclassified than the PACL group. Findings suggests that a low score might represent an increase in injury risk on a group level but does not predict a future injury on the individual level. To predict future cases of a multi-factorial injury like an ACL it is possible that a multi-factorial dataset (multiple exercises, anthropometric data and so on) / a complex system model (Bittencourt et al., 2016) is needed as information from different domains need to be interlinked. Also, adding other factors such as exposure / training load may improve the efficacy of these methods.

Comparing effect sizes observed during the average and maximal analysis to the range of effect sizes observed in the shuffle split analysis, suggests that there is an impact of utilizing the average value of a measure and the maximal performance trial of n-trails to an analysis. When considering the shuffle split as “truth”, the No Injury and PACL comparison overestimated the effect of the differences in the average analysis, while in the No Injury and FACL comparison the maximal comparisons seemed to have overestimated the effect of difference between the groups.

Previous ACL Injuries

The observed difference in the healthy score between the No Injury and PACL group was of medium effect (average analysis d = 0.53; maximal analysis = 0.45). As such, the classification model tested can be seen as relatively robust because the validation cohort differed from the original cohort, which was the basis for the movement model, in both sex (female vs male), test execution (vertical drop jump with arm swing) and partly in the type of sport (football and handball vs. rugby, football and Gaelic games). Considering that players with existing ACL injury will have a much higher risk (up to 15 times) for a new ACL injury (Paterno et al., 2014; Wiggins et al., 2016), it is possible that there are similar underlying movement characteristics that can predispose athletes for the first or second ACL injury. The classification model, which was purely data driven, identified asymmetries as key factors that separated the movement characteristics of ACL injured and healthy players. This suggests that movement characteristics displayed in a vertical drop jump might be associated with injury risk in players with or without a previous injury and that movement asymmetry seems to be relevant for injury risk, supporting what has been suggested in previous risk factor studies (Paterno, Rauh, Schmitt, Ford, & Hewett, 2012; Paterno et al., 2011; Webster & Gribble, 2010), and what is considered best practice for return to sport (Grindem, Snyder-Mackler, Moksnes, Engebretsen, & Risberg, 2016; King et al., 2019; Kyritsis, Bahr, Landreau, Miladi, & Witvrouw, 2016),

However, although significant differences were observed in healthy score between the healthy and PACL group in this validation study, the classification accuracy was much lower compared with the accuracy previously observed in the original cohort (AUC = 0.94 vs 0.53; Richter et al. [2019]). One explanation for the poor classification accuracy is that the cohort used for the Krosshaug et al. (2016) data contained only athletes that were fully rehabilitated and back in normal game play, whereas the cohort used to build the classification algorithm (rugby and football Gaelic games 9 months after ACL surgery) had just returned or were about to return to play.

Future ACL Injuries

When comparing the healthy score (max. jump) between the No Injury and FACL group, a significant difference with a small effect size (average analysis d = 0.28; maximal analysis = 0.36) was found between the two groups. The ability to identify individual athletes with increased risk for a new injury was poor, meaning that the developed classification algorithm cannot be used for to identify athletes who are at higher risk for sustaining future ACL injuries. The poor predictive ability of the algorithm for identifying new injuries is not surprising, considering that the algorithm was not trained for this purpose. The observed prediction ability is not much better than guessing, which is in line with previous studies that investigate the risk of sustaining injury in players without a previous injury (.51 < AUC < .70; Krosshaug et al., 2016; Leppänen et al., 2017; Smith et al., 2012). Although small a pilot study by Hewett et al. (2005) reported a high specificity (0.73) and sensitivity (0.78) using dynamic valgus angles or moments – these findings have not been replicated in later, much larger cohorts (Krosshaug et al., 2016; Leppänen et al., 2017).

Still, it should be noted that, in contrast to the traditional analysis approach of Krosshaug et al. [14], the classification model was able to detect small differences (on a group level) in the jump biomechanics between healthy players and those who went on to sustain an ACL injury. Further, it demonstrated an injury incidence proportion higher than the sample average (6.7%) for athletes with a healthy score below 0.33 ($p < 0.001$; 7.6%) and an injury incidence proportion lower than the sample average for athletes with a healthy score above 0.66 ($p < 0.001$; 5.5%).

This suggests that differences can be discovered by a framework that utilizes a mathematical approach to reduce waveforms and machine learning techniques (e.g., regression to neural networks) that may not be found using a classical discrete point analysis. It is also important to note, that the accuracy measures reported here are likely to underestimate the "true" ability of a classification algorithm, since it was not trained for this purpose.

Interestingly, the algorithm did not include features suggested to be important based on the current literature (e.g. peak knee abduction moment; Bates et al., 2018), nor did features capture the early movement cycle/impact phase, which correspond to the critical time period where ACL injuries occur (Koga et al., 2011, 2010). This makes the comparison to previous published injury studies difficult, highlights that the classification model should be reworked, but also highlights that alternative variables should be considered in future studies (as the classification model was able to find differences, which wasn't the case in previous analysis).

Classification models of future studies should seek to combine expert-knowledge-driven and data-driven features extraction and utilize a variety of movement task to enable the "learning" of the underlying function a multiple injury / injury pre-disposing pattern.

This study did not perform any adjustments for multiple comparisons. To estimate the impact of alpha 1 errors to findings, a shuffle split analysis was performed demonstrating the frequency of significant findings within random selections of the examined dataset.

Conclusions

We observed lower healthy scores in athletes with previous ACL injuries compared to healthy players in the validation cohort. Healthy scores were also significantly lower, on a group-level, in athletes who went on to sustain a future injury compared with uninjured athletes. This suggests that a low healthy score might represent an increase in injury risk on a group level but does not predict a future injury on the individual level. Utilized features indicate that limb symmetry and jump ability may be linked to increased ACL injury risk. However, the ability to identify elite female athletes with either an existing or a future injury was poor, meaning that the tested classification algorithm cannot be used for this purpose.

References

Barton, G. J., Hawken, M. B., Scott, M. A., & Schwartz, M. H. (2012). Movement Deviation Profile: A measure of distance from normality using a self-organizing neural network. *Human Movement Science*, *31*(2), 284–294.

Bates, N. A., Schilaty, N. D., Nagelli, C. V., Krych, A. J., & Hewett, T. E. (2018). Validation of Noncontact Anterior Cruciate Ligament Tears Produced by a Mechanical Impact Simulator Against the Clinical Presentation of Injury. *American Journal of Sports Medicine*, *46*(9), 2113–2121. https://doi.org/10.1177/0363546518776621

Bittencourt, N. F. N., Meeuwisse, W. H., Mendonça, L. D., Nettel-Aguirre, A., Ocarino, J. M., & Fonseca, S. T. (2016). Complex systems approach for sports injuries: Moving from risk factor identification to injury pattern recognition - Narrative review and new concept. *British Journal of Sports Medicine*, *50*(21), 1309–1314. https://doi.org/10.1136/bjsports-2015-095850

Borenstein, M., Hedges, L. V, Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction To Meta Aalysis*. John Wiley & Sons.

Chawla, N., Bowyer, K., Hall, L., & Kegelmeyer, W. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Cohen, J. (1988). *Statistical power analysis for the behavioral sciencies*. Routledge.

Grindem, H., Snyder-Mackler, L., Moksnes, H., Engebretsen, L., & Risberg, M. A. (2016). Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: The Delaware-Oslo ACL cohort study. *British Journal of Sports Medicine*, *50*(13), 804–808. https://doi.org/10.1136/bjsports-2016-096031

Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*. Springer US. https://doi.org/10.1038/s41591-018-0268-3

Hewett, T. E., Myer, G. D., Ford, K. R., Heidt, R. S., Colosimo, A. J., McLean, S. G., … Succop, P. (2005). Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes. *The American Journal of Sports Medicine*, *33*(4), 492–501.

Hosmer, D. W., Stanley, L., & Rodney, S. X. (2013). *Applied logistic regression* (398th ed.). John Wiley & Sons.

Kerby, D. S. (2014). The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation. *Comprehensive Psychology*, *3*, 11.IT.3.1. https://doi.org/10.2466/11.it.3.1

King, E., Richter, C., Franklyn-Miller, A., Wadey, R., Moran, R., & Strike, S. (2019). Back to Normal Symmetry? Biomechanical Variables Remain More Asymmetrical Than Normal During Jump and Change-of-Direction Testing 9 Months After Anterior Cruciate Ligament Reconstruction. *American Journal of Sports Medicine*, *47*(5), 1175–1185. https://doi.org/10.1177/0363546519830656

Koga, H., Bahr, R., Myklebust, G., Engebretsen, L., Grund, T., & Krosshaug, T. (2011). Estimating anterior tibial translation from model-based image-matching of a noncontact anterior cruciate ligament injury in professional football: A case report. *Clinical Journal of Sport Medicine*, *21*(3), 271–274. https://doi.org/10.1097/JSM.0b013e31821899ec

Koga, H., Nakamae, A., Yosuke, S., Iwasa, J., Myklebust, G., Engebretsen, L., … Krosshaug, T. (2010). Mechanisms for Noncontact Anterior Cruciate Ligament Injuries: Knee Joint Kinematics in 10 Injury Situations From Female Team Handball and Basketball. https://doi.org/10.1177/0363546510373570

Krosshaug, T., Steffen, K., Kristianslund, E., Nilstad, A., Mok, K.-M., Myklebust, G., … Bahr, R. (2016). The vertical drop jump is a poor screening test for ACL injuries in female elite soccer and handball players: a prospective cohort study of 710 athletes. *The American Journal of Sports Medicine*, *44*(4), 874–883.

Kyritsis, P., Bahr, R., Landreau, P., Miladi, R., & Witvrouw, E. (2016). Likelihood of ACL graft rupture: Not meeting six clinical discharge criteria before return to sport is associated with a four times greater risk of rupture. *British Journal of Sports Medicine*, *50*(15), 946–951. https://doi.org/10.1136/bjsports-2015-095908

Leppänen, M., Pasanen, K., Kujala, U. M., Vasankari, T., Kannus, P., Äyrämö, S., … Parkkari, J. (2017). Stiff Landings Are Associated with Increased ACL Injury Risk in Young Female Basketball and Floorball Players. *American Journal of Sports Medicine*, *45*(2), 386–393. https://doi.org/10.1177/0363546516665810

Paterno, M. V., Rauh, M. J., Schmitt, L. C., Ford, K. R., & Hewett, T. E. (2012). Incidence of contralateral and ipsilateral anterior cruciate ligament (ACL) injury after primary ACL reconstruction and return to sport. *Clin J Sport Med*, *22*(2), 116–121.

Paterno, Mark V., Rauh, M. J., Schmitt, L. C., Ford, K. R., & Hewett, T. E. (2014). Incidence of second ACL injuries 2 years after primary ACL reconstruction and return to sport. *American Journal of Sports Medicine*, *42*(7), 1567–1573. https://doi.org/10.1177/0363546514530088

Paterno, Mark V, Schmitt, L. C., Ford, K. R., Rauh, M. J., Myer, G. D., & Hewett, T. E. (2011). Effects of sex on compensatory landing strategies upon return to sport after anterior cruciate ligament reconstruction. *Journal of Orthopaedic and Sports Physical Therapy*, *41*(8), 553–559. https://doi.org/10.2519/jospt.2011.3591

Richter, C., King, E., Strike, S., & Franklyn-Miller, A. (2019). Objective classification and scoring of movement deficiencies in patients with anterior cruciate ligament reconstruction. *PLOS ONE*, *14*(7), e0206024. https://doi.org/10.1371/journal.pone.0206024

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, 43–46.

Smith, H. C., Johnson, R. J., Shultz, S. J., Tourville, T., Holterman, L. A., Slauterbeck, J., … Beynnon, B. D. (2012). A prospective evaluation of the Landing Error Scoring System (LESS) as a screening tool for anterior cruciate ligament injury risk. *American Journal of Sports Medicine*, *40*(3), 521–526. https://doi.org/10.1177/0363546511429776

Webster, K. A., & Gribble, P. A. (2010). Time to stabilization of anterior cruciate ligament-reconstructed versus healthy knees in National Collegiate Athletic Association Division I female athletes. *Journal of Athletic Training*, *45*(6), 580–585.

Wiggins, A. J., Grandhi, R. K., Schneider, D. K., Stanfield, D., Webster, K. E., & Myer, G. D. (2016). Risk of secondary injury in younger athletes after anterior cruciate ligament reconstruction: a systematic review and meta-analysis. *American Journal of Sports Medicine*, *44*(7), 1861–1876.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3