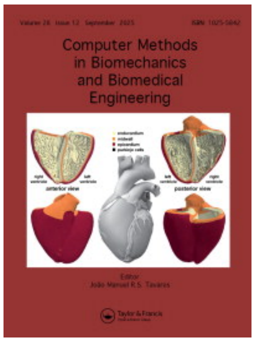

Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering

ISSN: 1025-5842 (Print) 1476-8259 (Online) Journal homepage: www.tandfonline.com/journals/gcmb20

Investigating ACL length, strain and tensile force in high impact and daily activities through machine learning

Elisa Roldán Ciudad , Neil D. Reeves , Glen Cooper & Kirstie Andrews

To cite this article: Elisa Roldán Ciudad , Neil D. Reeves , Glen Cooper & Kirstie Andrews (30 Aug 2025): Investigating ACL length, strain and tensile force in high impact and daily activities through machine learning, Computer Methods in Biomechanics and Biomedical Engineering, DOI: 10.1080/10255842.2025.2551846

To link to this article: https://doi.org/10.1080/10255842.2025.2551846

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING
https://doi.org/10.1080/10255842.2025.2551846

Investigating ACL length, strain and tensile force in high impact and daily activities through machine learning

Elisa Roldán Ciudad[a], Neil D. Reeves[b], Glen Cooper[c] and Kirstie Andrews[a]

[a]Department of Engineering, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK; [b]Lancaster Medical School, Faculty of Health and Medicine, Lancaster University, Lancaster, UK; [c]School of Engineering, University of Manchester, Manchester, UK

ABSTRACT

Anterior cruciate ligament (ACL) reconstruction rates are rising, particularly among female athletes, though causes remain unclear. This study: (i) identify accurate machine learning models to predict ACL length, strain, and force during six high-impact and daily activities; (ii) assess the significance of kinematic and constitutional parameters; and (iii) analyse gender-based injury risk patterns. Using 9,375 observations per variable, 42 models were trained. Cubist, Generalized Boosted Models (GBM), and Random Forest (RF) achieved the best R^2, RMSE, and MAE. Knee flexion and external rotation strongly predicted ACL strain and force. Female athletes showed higher rotation during cuts, elevating ACL strain and risk.

ARTICLE HISTORY
Received 21 November 2024
Accepted 18 August 2025

KEYWORDS
ACL biomechanics; anterior cruciate ligament (ACL); daily and high-impact activities; motion capture; knee model; machine learning models

Introduction

The anterior cruciate ligament (ACL) is one of the most frequently injured ligamentous tissues (Dargel et al. 2007). In the United States, over 200,000 ACL injuries are diagnosed annually, with around 175,000 requiring surgical intervention, costing an estimated \$1 billion per year (Leong et al. 2014). ACL reconstruction rates have increased by 37% in recent years, with an even greater rise of 47% observed among females (Lyman et al. 2009; Buller et al. 2015).

The ACL plays a crucial role in stabilising the knee by limiting anterior tibial translation and excessive external rotation (Andriacchi and Dyrby 2005). Sports such as football, basketball, volleyball, and skiing, which involve sudden decelerations, jumping, pivoting, and crossover cutting, account for approximately 78% of ACL injuries (Fleming et al. 1998; Prodromos et al. 2007). These injuries predominantly affect young, active individuals, with female athletes experiencing ACL injuries at a rate three to six times higher than their male counterparts (Agel et al. 2005; Prodromos et al. 2007). However, the reasons behind this gender disparity and the specific knee joint kinematics associated with higher injury risks in males and females remain unclear.

Identifying key kinematic and individual characteristics linked to ACL injury risk can improve understanding of injury mechanisms and gender-related differences. This knowledge can support injury prevention, enhance rehabilitation strategies, and enable the design of personalised ligament implants based on a patient's anatomy, movement patterns, and clinical needs (Roldán et al. 2016). Previous studies have investigated *in-vivo* ACL biomechanics during various activities (Roldán et al. 2016, 2017), and how these insights can guide implant design (Roldán et al. 2024a). This study advances the field by analysing critical factors influencing *in-vivo* ACL mechanics and identifying gender-specific risk patterns using a novel approach that integrates machine learning (ML) and interactive graphical analysis.

Both traditional statistical and ML models are effective for examining how kinematic and participant features affect ACL strain and force, helping to identify factors contributing to increased injury risk. Developing optimised, reproducible, and accurate models, with high coefficients of determination (R^2)

CONTACT Elisa Roldán Ciudad Elisa.Roldan-Ciudad@mmu.ac.uk

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10255842.2025.2551846.

and low error rates, reduces the need for extensive experiments, saving time and costs (Roldán et al. 2023b). ML techniques are particularly valuable, offering accurate predictions despite complex nonlinearities or unmet parametric conditions (Bzdok et al. 2018).

The use of machine learning (ML) in healthcare is rapidly expanding (Al Kuwaiti et al. 2023), proving valuable in areas such as medical imaging and diagnostics (Bedi et al. 2015; Esteva et al. 2017; Gudigar et al. 2021), patient care (Baig et al. 2017), medical research (Weissler et al. 2021), tissue engineering (Roldán et al. 2023a,b), drug delivery (Mak and Pichika 2019), wound dressings (Roldán et al. 2025), rehabilitation (Yang et al. 2018), and human motion analysis (Xiang et al. 2022).

A Scopus search for 'anterior cruciate ligament' AND 'machine learning' returned 76 articles. Most focused on predicting ACL reconstruction revision (Martin et al. 2022; Ye et al. 2022) or diagnosing ACL injuries using imaging and deep learning (Bien et al. 2018; Chang et al. 2019; Germann et al. 2020; Fritz et al. 2023). Only two studies addressed knee biomechanics in healthy individuals. Tedesco et al. used motion sensors with various ML algorithms—including KNN, naïve Bayes, SVM, gradient boosting, and multilayer perceptrons—to analyse gait in healthy and ACL-reconstructed rugby players (Tedesco et al. 2020). Chaaban et al. employed inertial sensors and stepwise linear regression to predict ground reaction forces and knee biomechanics during a double-limb jump landing (Chaaban et al. 2021). However, none of these studies aimed to accurately predict *in-vivo* ACL length, strain, or force in healthy participants across a wide range of daily and high-impact activities, nor did they investigate gender-based differences in ACL kinematics and dynamics.

This study had three main aims: (i) to identify the most accurate regression models for predicting ACL length, strain, and force using seven features: activity, sex, height, weight, knee flexion, external rotation, and abduction angles; (ii) to assess the influence of these variables on ACL biomechanics; and (iii) to explore gender differences in ACL force normalised to body weight (force/BW) and associated risk factors across six daily and high-impact activities: walking, running, jumping, one-leg jumping, sidestep cutting, and crossover cutting.

Using biomechanical data from 12 participants, ACL kinematics and dynamics were predicted through 42 machine learning models, including Cubist and EARTH regressors, which had not previously been applied in biomechanical contexts. The most accurate models were selected to evaluate the relevance of the independent variables. Interactive graphical analysis revealed gender-specific differences in ACL force/BW across activities, aiming to clarify why female athletes experience higher injury rates.

We hypothesised that: (1) novel, high-accuracy ML models would highlight the key features influencing ACL biomechanics and, when combined with visual analysis, offer better insights into injury mechanisms; and (2) women would exhibit greater ACL strain and force, potentially explaining their higher injury risk. These findings are expected to support improved injury prevention, rehabilitation strategies, and the development of personalised ACL implants.

Methods

Data collection

Participants

Twelve healthy young adults (7 males and 5 females; mean ± SD: age 27.3 ± 3.3 years, height 1.70 ± 0.09 m, mass 71.6 ± 15.5 kg) participated in this study. All procedures complied with the Declaration of Helsinki. Ethical approval was granted by the Manchester Metropolitan University Ethics Committee (Approval Number: SE141530), and informed written consent was obtained from all participants. Prior to data collection, participants completed the Knee Injury and Osteoarthritis Outcome Score (KOOS) and Hip Injury and Osteoarthritis Outcome Score (HOOS) to confirm the absence of any previous knee or hip injuries.

Protocol

Participant motion during six activities—walking, running, crossover cutting, sidestep cutting, vertical jumping with both legs, and single-leg horizontal jumping—was recorded using a 10-camera motion capture system (Vicon 612, Oxford Metrics, UK). Thirty-three reflective markers were placed on key anatomical landmarks of the upper and lower body (Figure 1). Activities were performed at self-selected speeds, and kinematic data were sampled at 100 Hz over three successful trials per participant. A detailed protocol was previously described by Roldán et al. (2017).

Kinematic data were processed using Vicon Nexus 1.8.5 (Vicon Motion Systems Ltd., UK) and analysed in OpenSim 3.3 (SimTK, Stanford, CA).

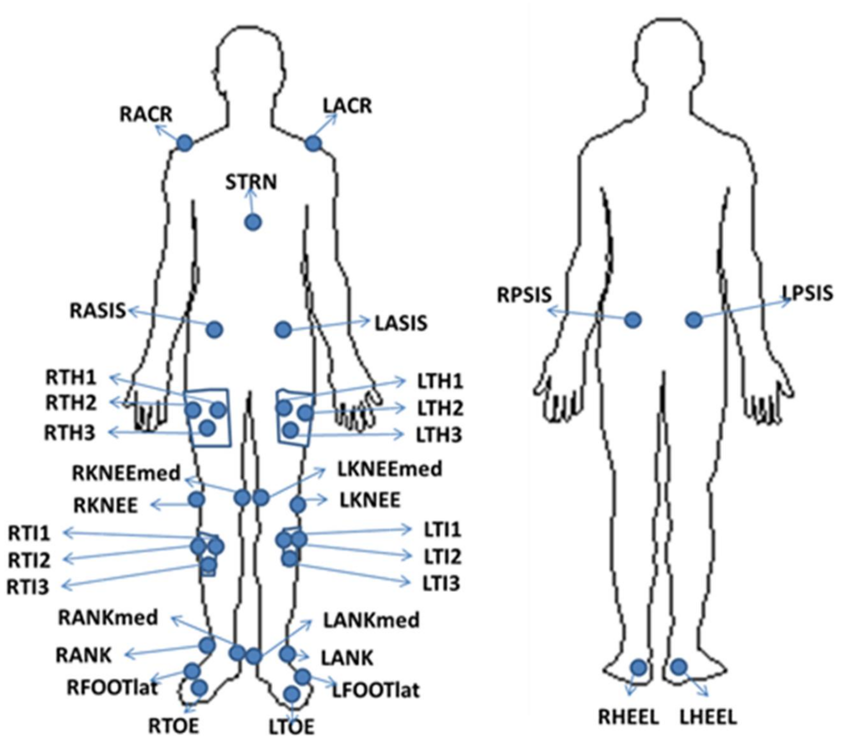

Figure 1. Reflective marker's locations.

ACL length, strain and force estimation

A 27-degree-of-freedom (DOF) OpenSim model, including 3 DOF per knee, 12 bones, and 92 musculotendon actuators, was created and scaled for each participant to estimate *in-vivo* right ACL length at each timepoint by tracking ACL insertion coordinates from processed kinematic data, as described by Roldán et al. (2017). Strain was calculated at every 0.01 s using Eqs. (1) and (2):

$$\varepsilon = \frac{L - L_0}{L_0} \tag{1}$$

$$L_0 = \frac{L_r}{(\varepsilon_r + 1)} \tag{2}$$

The zero-load length (L_0) was calculated using the ACL length at full knee extension (L_r, 0.024–0.035 m depending on participant geometry) and a reference strain (ε_r) of 0.08, based on (Blankevoort and Huiskes 1991).

ACL tensile forces were estimated using their established force–strain relationship, modelling the ACL as a non-linear passive elastic tissue (Roldán et al. 2016). To compare between individuals, forces were normalised to body weight (BW) following (Schmitt et al. 2015). Further details on participant-specific scaling, and the estimation of ACL length, strain, and force, are provided in Roldán et al. (2024a).

Data set

Following data collection, 9,375 observations were obtained for each independent variable (activity, height, weight, sex, knee flexion, external rotation, and abduction angles) and each dependent variable (ACL length, strain, and force/BW). These corresponded to the kinematic and dynamic data recorded at each timepoint during three successful repetitions of each activity per participant.

Initial statistical analysis

An initial exploratory analysis was conducted to examine variable distributions and correlations. Normality and homoscedasticity were tested using the Kolmogorov–Smirnov and Breusch–Pagan tests, with results provided in the supplementary material. All initial statistical analyses and machine learning models were performed using R 4.3.0 and RStudio 2023.03.1. Descriptive statistics comparing ACL force/BW between genders were carried out in IBM SPSS v.27 (IBM Inc., US).

Prediction models

Seven features (activity, height, weight, sex, knee flexion angle, knee external rotation angle, and knee abduction angle) were selected to predict three endogenous variables (ACL length, strain, and ACL force/BW). These variables were chosen to simplify the models and due to their reported influence on ACL loading in previous studies (Yoo et al. 2010; Taylor et al. 2013; Utturkar et al. 2013). 9,375 observations per feature were used to train 14 ML models: Linear Models (LM), Generalised Linear Model (GLM), Generalised Additive Model (GAM), Stepwise Model Selection by AIC (Step AIC), Multivariate Adaptive Regression Splines (EARTH_1), optimised EARTH (EARTH_2), GLMNET, Support Vector Machine (SVM), Classification and Regression Trees (CART_1), optimised CART (CART_2), k-Nearest Neighbours (KNN), Random Forest (RF), Generalised Boosted Models (GBM), and Cubist (CUBITS). This study is the first to apply Cubist and EARTH models in biomechanical research. A total of 42 ML models (14 per outcome variable) were analysed, each employing distinct prediction strategies—ranging from parametric regressions to rule-based and distance-based algorithms. Table 1 outlines model characteristics, with further details available in the supplementary material.

Classification and Regression Trees (CART) were included in this study due to their interpretability and ability to visually represent prediction rules and feature importance. Two CART models were developed to predict ACL length, strain, and force/BW and assess the influence of the seven independent variables (features). The first model (CART_1) was created using the caret() library with the 'rpart' method and hyperparameters of 0, 0.05, and 0.1. The second model (CART_2) used the rpart() and rpart.plot() libraries, with the 'anova' method and the prune() function to identify the most influential variables.

All ML models were developed using the caret() library, as recommended in previous studies (Roldán et al. 2024b,c), due to its integrated functions for data preprocessing, hyperparameter tuning, and model training. Each model was trained using its corresponding method: 'svmRadial' for SVM, 'rpart' for CART, and 'lm', 'glm', 'earth', 'gam', 'glmStepAIC', 'glmnet', 'knn', 'rf', 'gbm', and 'cubist' for the respective approaches. For the optimised EARTH models (EARTH_2), hyperparameters were tuned by varying the degree (set to 1) and the number of prunes (2, 11, and 10).

Table 1. Characteristics of the ML models used in the present study and library used to build each model.

Models	Parametric	Regression[1]	Rules	Distances	Adaptative	Library
LM	YES	YES	NO	NO	NO	caret
GLM	YES	YES	NO	NO	YES	caret
EARTH_1	MIX	YES	YES	NO	YES	caret
EARTH_2	MIX	YES	YES	NO	YES	earth
GAM	MIX	YES	NO	NO	YES	caret
GLMNET	YES	YES	NO	NO	YES	caret
StepAIC	YES	YES	NO	NO	YES	caret
CART_1	NO	NO	YES	NO	NO	caret
CART_2	NO	NO	YES	NO	NO	rpart
KNN	NO	NO	NO	YES	NO	caret
SMV	NO	NO	NO	YES	NO	caret
RF	NO	NO	YES	NO	NO	caret
GBM	NO	NO	YES	NO	NO	caret
Cubist	NO	YES	YES	YES	NO	caret

([1]) All models are for regression and classification. However, some of them do not use the regression in their algorithms.

Preprocessing was standardised across models using 'center' and 'scale', except for RF, GBM, and CUBIST, which employed the 'BoxCox' transformation. All models were evaluated using nested cross-validation (Figure 2). The outer loop applied a 12-fold leave-one-participant-out approach to ensure full independence between training and test sets and avoid data leakage. Each outer loop's training set included data from 11 participants, while the test set included data from the remaining participant. Inner loops 11-fold cross-validation (3 repeats) were created with the 'trainControl()' function, 'repeatedcv' method, and 'expand.grid()' to optimise hyperparameters for EARTH and CART models. Default settings in caret() were used for the rest of the models. To prevent library conflicts, the tidymodels package was used.

To identify the most accurate model for each dependent variable, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2) were computed using the functions 'RMSE()', 'MAE()' and 'R2()' by comparing the predicted values (generated with the predict() function) against the observed data.

Final cross-validation metrics (R^2_{CV}, MAE_{CV} and $RMSE_{CV}$) were calculated as the average of individual participant results (R^2_{CV}_P1 to R^2_{CV}_P12, MAE_{CV}_P1 to MAE_{CV}_P12, and $RMSE_{CV}$_P1 to $RMSE_{CV}$_P12), which were calculated from the average values across each inner validation fold (R^2_{CV1} to R^2_{CV11}, MAE_{CV1} to MAE_{CV11} and $RMSE_{CV1}$ to $RMSE_{CV11}$) based on training and validation data (as illustrated in Figure 2).

To determine the most accurate model per independent variable, the final MAE and RMSE errors (MAE_{Test} and $RMSE_{Test}$), and R^2 (R^2_{Test}) were computed by averaging the test results across all participants (R^2_{Test1} to R^2_{Test12}, MAE_{Test1} to MAE_{Test12} and $RMSE_{Test1}$ to $RMSE_{Test12}$) as shown in Figure 2.

Importance of the independent variables on the dependent variables

For all ML regression models, the importance of the features in predicting the endogenous variables was assessed using the 'varImp()' function from the 'caret()' library, applied to the training and validation datasets.

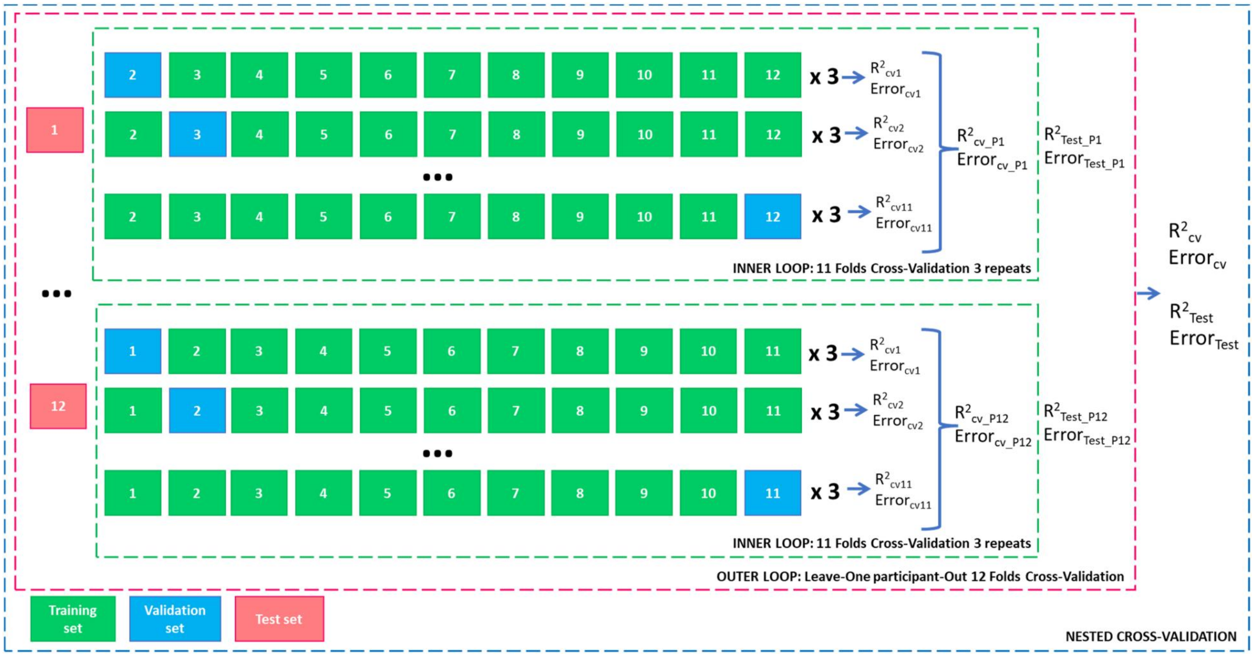

Figure 2. Validation of the models.

The percentual average of importance of each feature was calculated by considering all ML models (14 models per dependent variable), all participants (12 in total), and all cross-validation repetitions (33 per participant). Consequently, the importance of each predictor with respect to each dependent variable (ACL length, ACL strain, and ACL force/BW) was determined from a total of 5,544 cases per outcome.

Identification of gender differences for ACL force/BW

The relationship between ACL force/BW versus the ACL strain was explored for each activity and gender to determine which activities presented the highest ACL force for each gender. All the ACL force/BW and ACL strain observations (9375 observations each) were included in this analysis.

Percentiles 25, 50, 75, 95 and 99 for all the independent and dependent variables were calculated for each gender and activity to identify thresholds that revealed gender-based discrepancies.

After identifying values above the 75th percentile as those showing the greatest gender differences, ACL force/BW and strain data above this threshold were analysed to examine the associated knee angles and explore potential biomechanical reasons for the higher incidence of ACL injuries in female athletes compared to males.

The libraries 'ggplot2()', 'scatterplot3D()' and 'plotly()' were used to generate visualisations for male and female participants and for each activity.

The full methodology is summarised in Figure 3.

Results and discussion

Initial statistical analysis

Exploratory analysis was used to assess the distribution of endogenous and exogenous variables (mean, variance, skewness, and kurtosis) and guide the selection of appropriate analytical methods. The three joint angles (flexion, external rotation, abduction) and two response variables (ACL length and strain) showed quasi-symmetric platykurtic distributions ($k < 3$) with similar mean and median values. In contrast, ACL force/BW followed a Pareto distribution, with 70–75% of values near zero and the remainder increasing to a peak of 3.04 N/BW, recorded in a female during crossover cutting, with knee flexion and external rotation slightly above 100° and 25°, respectively. Normality and homoscedasticity tests (Kolmogorov–Smirnov and Breusch–Pagan) confirmed that none of the three output variables were normally distributed or homoscedastic ($p < 0.001$), indicating that linear regression models were unsuitable, although they were included for comparison in line with previous studies (Roldán et al. 2023b). Full results are provided in the supplementary material.

Machine learning model selection

The models included seven input variables, two discrete (activity, sex) and five continuous (height, weight, knee flexion, external rotation, and abduction angles). As the data did not meet parametric assumptions of normality,

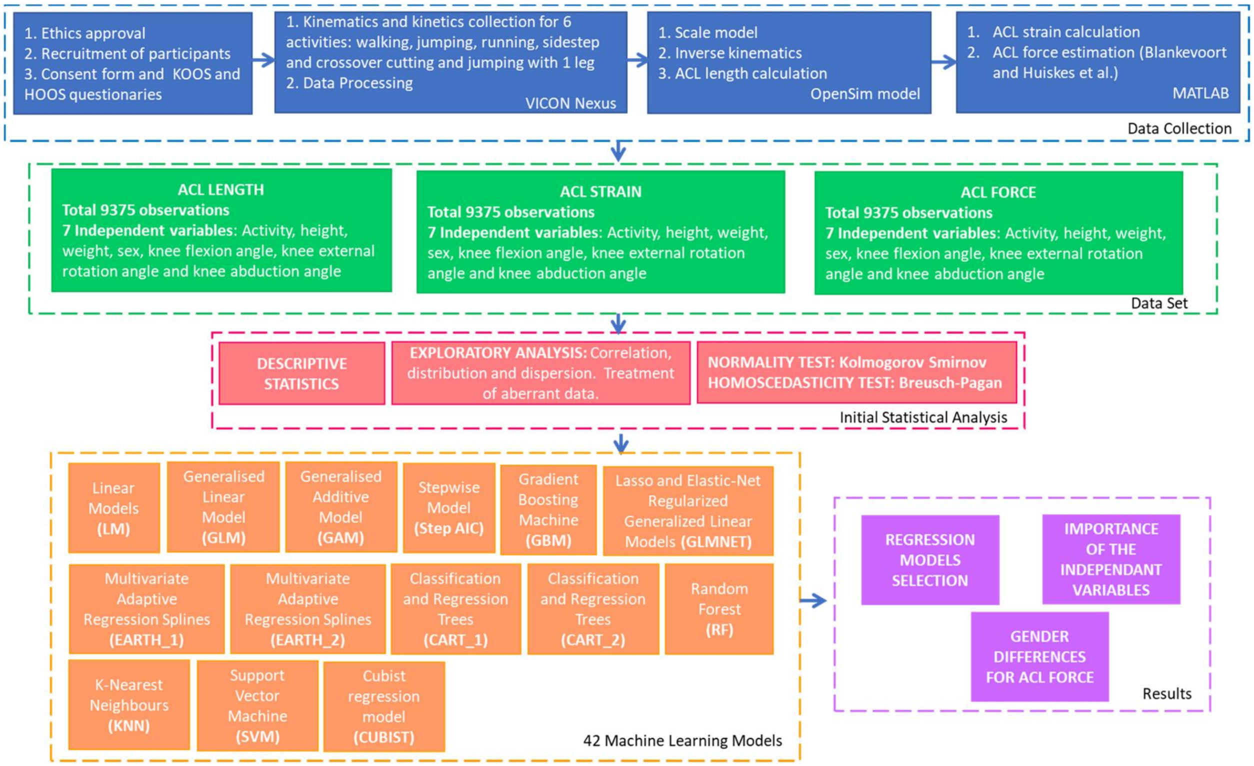

Figure 3. Outline of the followed methodology from human biomechanical data collection through to application of machine learning algorithms.

non-parametric machine learning models were deemed more appropriate for predicting the endogenous variables: ACL length, ACL strain, and ACL force/BW. While the coefficient of determination (R^2) reflects model fit, it does not alone indicate model quality; therefore, R^2 was evaluated alongside error metrics (MAE and RMSE) to determine model accuracy. Graphical representations of these metrics were used to identify the most precise models. Figure 4 presents R^2, RMSE, and MAE for each model predicting ACL length. Specifically, Figure 4A displays R^2 values, Figure 4B shows RMSE, and Figure 4C reports MAE, calculated through cross-validation and test data as detailed in the methods.

All ML models predicting ACL length, except GLM, LM, GLMNET and Step AIC, showed high R^2 values (0.987–0.996) and low RMSE (0.000521–0.000166) and MAE (0.00029–0.00011 m) after cross-validation. Test data revealed that RF, GBM and CUBIST achieved RMSE and MAE below 0.002, while EARTH, GBM and GAM yielded the highest R^2 (0.957–0.984). GBM was the most accurate model, with $R^2 = 0.991$, RMSE $= 0.000338$ and MAE $=$ 0.000256 m in cross-validation, and $R^2 = 0.961$, RMSE $= 0.00194$ and MAE $= 0.00188$ m with test data. Figure 5 presents model performance for ACL strain prediction, with panels Figure 5A–5C showing R^2, RMSE and MAE.

CUBIST, RF and GBM achieved the best fit for ACL strain prediction, with cross-validation R^2 between 0.971 and 0.995, and test data R^2 between 0.744 and 0.775. CUBIST was the most accurate overall, with $R^2 = 0.995$, RMSE $= 0.0059$ and MAE $=$ 0.0036 during cross-validation, and $R^2 = 0.744$, RMSE $= 0.196$ and MAE $= 0.111$ with test data, showing the lowest errors despite a slightly lower R^2 than RF (0.761) and GBM (0.775). Figure 6 summarises the performance of each model for ACL force/BW prediction, with panels Figure 6A–6C showing R^2, RMSE and MAE.

As with ACL strain, GBM, RF and CUBIST were the most accurate models for ACL force/BW prediction, with cross-validation R^2 of 0.947, 0.987 and 0.993 N/BW, and test R^2 of 0.775, 0.761 and 0.744 N/BW, respectively. RMSE and MAE were below 0.074 and 0.045 in cross-validation, and under 0.211 and 0.127 in test data. Despite slightly lower R^2, CUBIST had the lowest errors (RMSE $= 0.196$; MAE $=$ 0.111 N/BW) and was selected as the best model. As expected, parametric models such as GLM and LM underperformed compared to non-parametric ML models, as shown in prior studies (Rahbar and Vadood 2015; Kalantary et al. 2020; Roldán et al. 2023b, 2024b, 2024c). RF outperformed CART in all outputs due to its ensemble approach improving

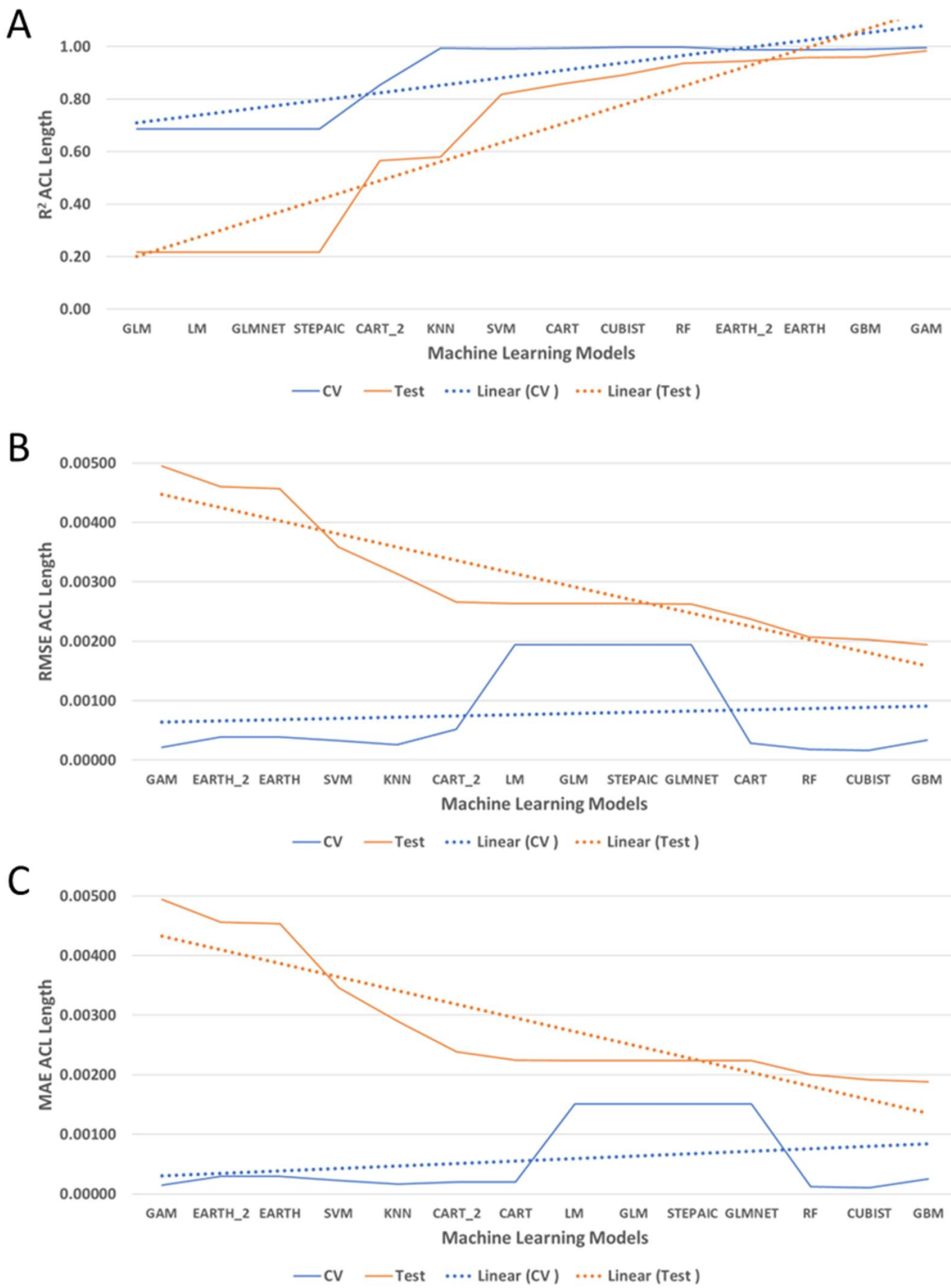

Figure 4. Evaluation of ML models to predict ACL length A) R^2 calculated from test data and with cross-validation, B) RMSE calculated from test data and with cross-validation, C) MAE (m) calculated from test data and with cross-validation.

prediction stability (Muqeet et al. 2023). Models built with specific libraries ('earth', 'rpart') yielded slightly better performance than those from the 'caret()' package, which automates fitting but reduces control. Nonetheless, 'caret()' was favoured for efficiency, ease of use, and computational speed. Figures 4–6 show that rule-based models (GBM, RF, CUBIST) achieved the highest R^2 and lowest errors, as they effectively handle non-linearity, ensemble learning, and overfitting. In contrast, regression-based models (GAM, StepAIC, LM, GLM, GLMNET, EARTH) yielded poorer performance, consistent with the non-linear nature of ACL biomechanics (see Figure 8C–D). The high predictive accuracy, especially of CUBIST with test data, confirms model reproducibility, validates variable contributions, and enhances understanding of ACL biomechanics. Further details are available in the supplementary material.

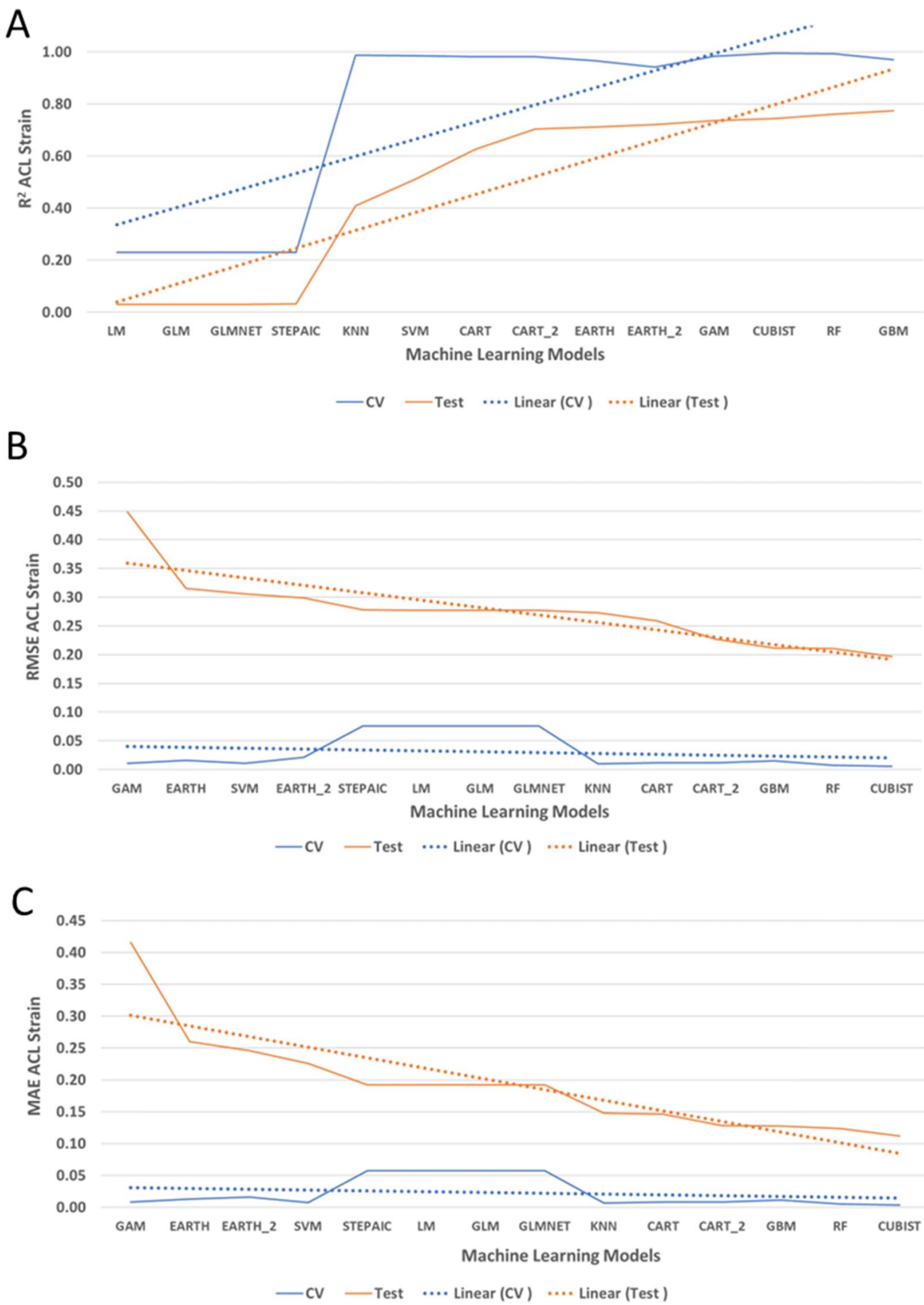

Figure 5. Evaluation of ML models to predict ACL strain A) R^2 calculated from test data and with cross-validation, B) RMSE calculated from test data and with cross-validation, C) MAE calculated from test data and with cross-validation.

Importance of the independent variables on the dependent variables

The influence of each independent variable on the three dependent variables was analysed across all ML models using 5544 cases per variable. Average percentage contributions are shown in Table 2.

As expected, ACL length was primarily influenced by participant constitution, especially height, taller individuals had longer unloaded and *in-vivo* ACLs throughout activity. Knee flexion angle had the highest impact on ACL length, consistent with previous studies (Yoo et al. 2010; Taylor et al. 2013; Roldán et al. 2017; Kono et al. 2020), and

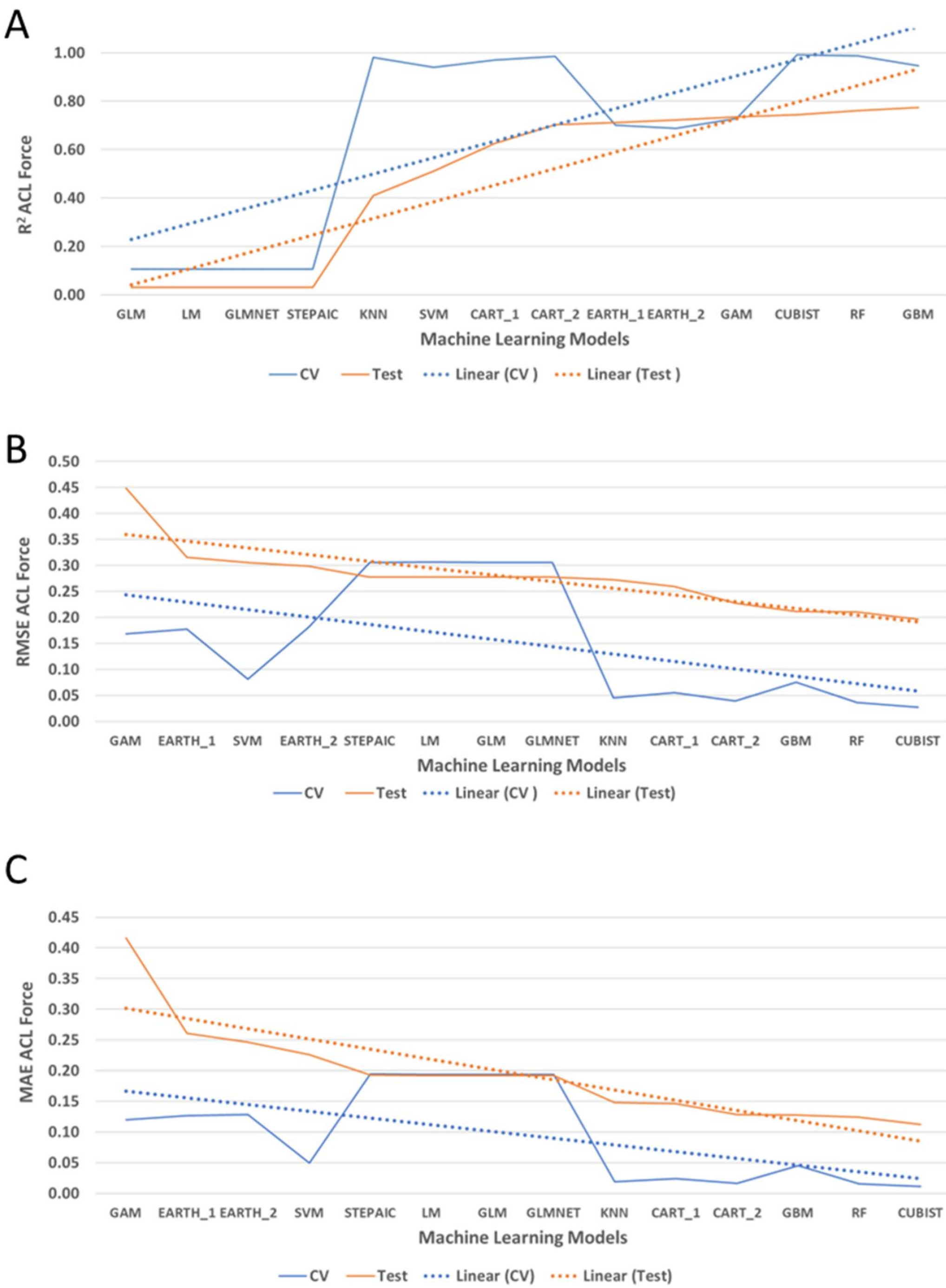

Figure 6. Evaluation of ML models to predict ACL force/BW A) R^2 calculated from test data and with cross-validation, B) RMSE calculated from test data and with cross-validation, C) MAE (N/BW) calculated from test data and with cross-validation.

knee external rotation was also significant (11.99%), confirming prior findings (Roldán et al. 2017). For ACL strain, knee flexion (34.69%) and external rotation (18.44%) were the most influential, followed by activity (11.09%), aligning with (Roldán et al. 2017), who identified two-leg jumps at maximum effort as inducing the greatest strain. As strain is a normalised measure, participant height had minimal influence. Similarly, ACL force/BW was most affected by knee flexion (33.16%) and external rotation (16.14%), followed by activity (13.98%). The two-leg jump produced the highest average ACL force/BW (1.076 ± 0.113 N/BW), as in Roldán et al. (2017), however the peak value (3.04 N/BW) occurred during crossover cutting in females. ACL force/BW was normalised following (Schmitt et al. 2015), thus height and weight had limited impact on its prediction.

CART models are effective for assessing the statistical contribution of independent variables to dependent outcomes due to their simplicity, support for both parametric and non-parametric data, robustness to outliers, and ability to handle skewed data without transformation (Song and Lu 2015; Roldán et al. 2023b).

Analysis of the pruned CART predicting ACL force/BW revealed that only 3.2% of observations involved knee flexion >88.2°. Within this subset, if knee rotation exceeded 14.1° (1.1%), ACL force/BW averaged 1.84 N/BW; if rotation was <14.1° (2.1%), it decreased to 0.672 N/BW. Hyperextension >2.41° (3.2%) resulted in an average force of 0.735 N/BW. These results align with prior studies linking high flexion with high rotation or hyperextension to elevated ACL forces and injury risk (Shimokochi and Shultz 2008; Quatman and Hewett 2009). For knee flexion between 9.52° and −2.41° with rotation >22.8% (1.6%), force averaged 0.746 N/BW, compared to 0.243 N/BW when rotation was <22.8° (14.5%). The lowest ACL force (0.05 N/BW) occurred with flexion between 88.2° and 9.52°, observed in 77.5% of cases, consistent with literature showing that 15–60° of flexion reduces ACL force (Mesfar and Shirazi-Adl 2006; Quatman and Hewett 2009). Figure 7 presents the pruned CART model. An optimised participant-specific version, including knee abduction contribution (7.19%), is provided in the supplementary material. Due to its complexity and low importance, knee abduction is not included in the main text.

Exploring gender differences for the ACL force in daily and high impact activities

ACL strain analysis revealed no significant sex differences below the 75th percentile; however, women exhibited 3.62 times higher strain than men at the 75th percentile. ACL force/BW was similar across sexes below the 50th percentile but increased notably in women thereafter: 4.65 times higher at the 50th, 11.3 times at the 75th, and 2.25 times at the 99th percentile. Among 9375 observations, peak ACL force/BW in males was 1.33 N/BW during the flight phase of a maximal jump, while in females it was 3.04 N/BW during crossover cutting, consistent with prior findings (Roldán et al. 2016, 2017; Englander et al. 2019; Foody et al. 2023). These results support existing evidence of higher ACL loading and injury risk in women, who suffer ACL injuries 3–6 times more than men (Agel et al. 2005; Prodromos et al. 2007).

At the 50th percentile, women showed 1.29 times greater flexion, 2.36 times greater external rotation, and men had 2.02 times greater abduction. Median ACL length was 18% longer in men. All relevant data, percentiles, and graphs are in the supplementary material.

The most significant gender differences in ACL strain and force/BW occurred above the 75th percentile, particularly with flexion >100°, rotation >25°, or hyperextension. Filtering by these thresholds revealed that during running, sidestep, and crossover cutting, women exhibited 20% more flexion and nearly triple the external rotation, leading to higher ACL force/BW and injury risk. Male peak values occurred during jumping: 1.2 N/BW during landing (120° flexion, <2° rotation) and 1.33 N/BW during flight with >5° hyperextension. These represent the highest point values across all male observations; average peak forces by activity were previously reported (Roldán et al. 2017).

A graphical analysis (Figure 8A) showed ACL force/BW vs strain, with a toe region followed by a linear region typical of soft tissues (Sharabi 2022). Women exhibited higher force peaks in all activities, that could lead to risk of fatigue-induced ACL failure, especially in running, sidestep, and crossover cutting, where women exceeded 2 N/BW and men stayed below 0.5 N/BW. Differences were smaller during walking and jumping, with women reaching ∼1.25 times higher force.

Flexion was the most influential variable on ACL force/BW. Figure 8C shows force/BW by activity and flexion angle, confirming that forces increase near full extension/hyperextension and flexion >100°, supporting prior findings that ACL length is minimised between 20 and 90° flexion (Kono et al. 2020). Women showed higher forces at flexion >80°.

Figure 8D (ACL force/BW vs flexion and rotation) showed women had greater rotation at flexion >100°, leading to increased ACL length, strain, and force. This aligns with previous reports showing females are more prone to ACL injury under flexion, high rotation, and valgus, while males are more vulnerable with high flexion and low rotation (Quatman and Hewett 2009).

Finally, Figure 8B presents predicted vs observed ACL force/BW using the CUBIST model, the most accurate model. Its reproducibility confirms that both predicted and observed data are equally valid for analysis.

To better identify scenarios associated with the highest ACL strain and force, all 9375 observations of

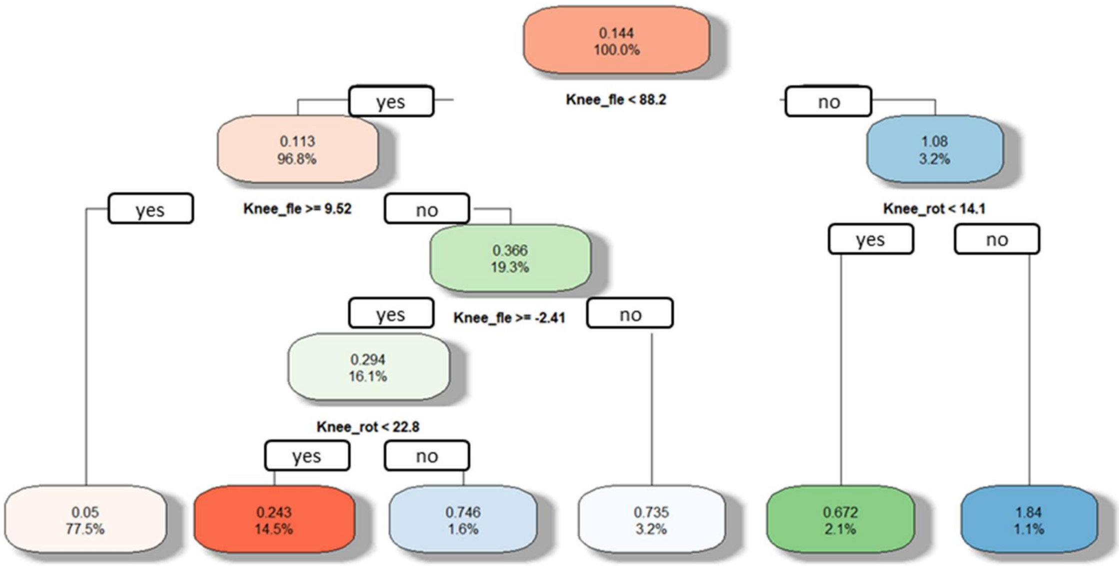

Figure 7. Pruned CART for ACL force/BW (N/BW).

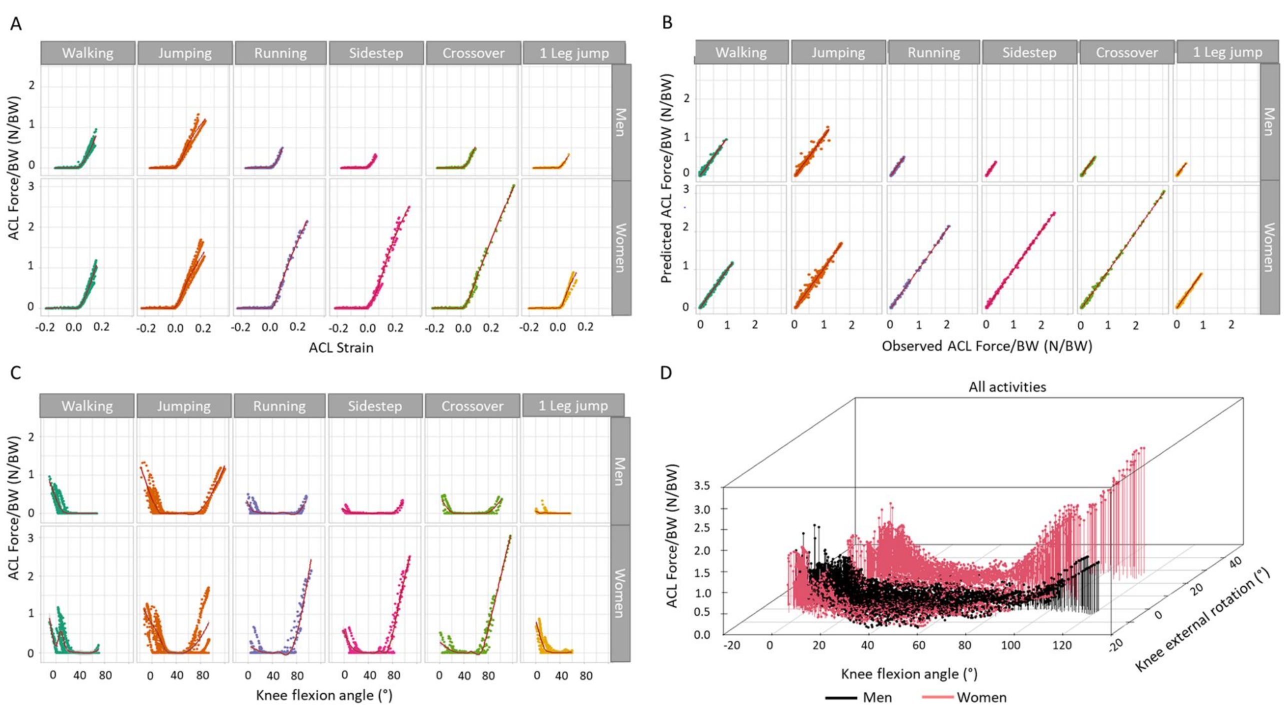

Figure 8. A) ACL force/BW vs ACL strain by activity and sex, B) predicted ACL force/BW from the CUBIST model vs observed ACL force/BW, C) ACL force/BW vs knee flexion by activity and sex and D) ACL force/BW vs knee flexion vs knee rotation by sex.

ACL strain and force/BW were analysed graphically against the three knee angles, and sex using interactive plots. Figures 9A and 9B show that ACL strain exceeded 10% during deep flexion (>100°) and near full extension or hyperextension in both sexes. Figures 9C and 9D indicate that women exhibited greater rotation and lower varus than men under high flexion, resulting in higher ACL force/BW and potentially a greater risk of injury due to long-term fatigue. Full interactive plots stratified by activity and sex are available in the supplementary material, offering detailed insights into the riskiest movements for each gender. Figure 9 presents a simplified overview of these analyses.

Limitations of the study

The 3DoF knee model used in this study assumes negligible translations compared to rotations in physiological conditions. Future studies should

Table 2. Importance of the independent variables on the ACL length, strain and force/BW.

ACL length		ACL strain		ACL force/BW	
Variables	Importance (%)	Variables	Importance (%)	Variables	Importance (%)
Heigh	32.96	**Knee_flexion_r**	34.69	**Knee_flexion_r**	33.16
Knee_flexion_r	26.28	**Knee_rotation_r**	18.44	**Knee_rotation_r**	16.42
Weight	13.98	**Activity**	11.09	**Activity**	13.98
Knee_rotation_r	11.99	**Heigh**	10.84	**Weight**	10.96
Sex	5.76	**Weight**	10.08	**Knee_abduction_r**	9.71
Knee_abduction_r	5.65	**Knee_abduction_r**	9.68	**Sex**	9.06
Activity	3.37	**Sex**	5.18	**Heigh**	6.70

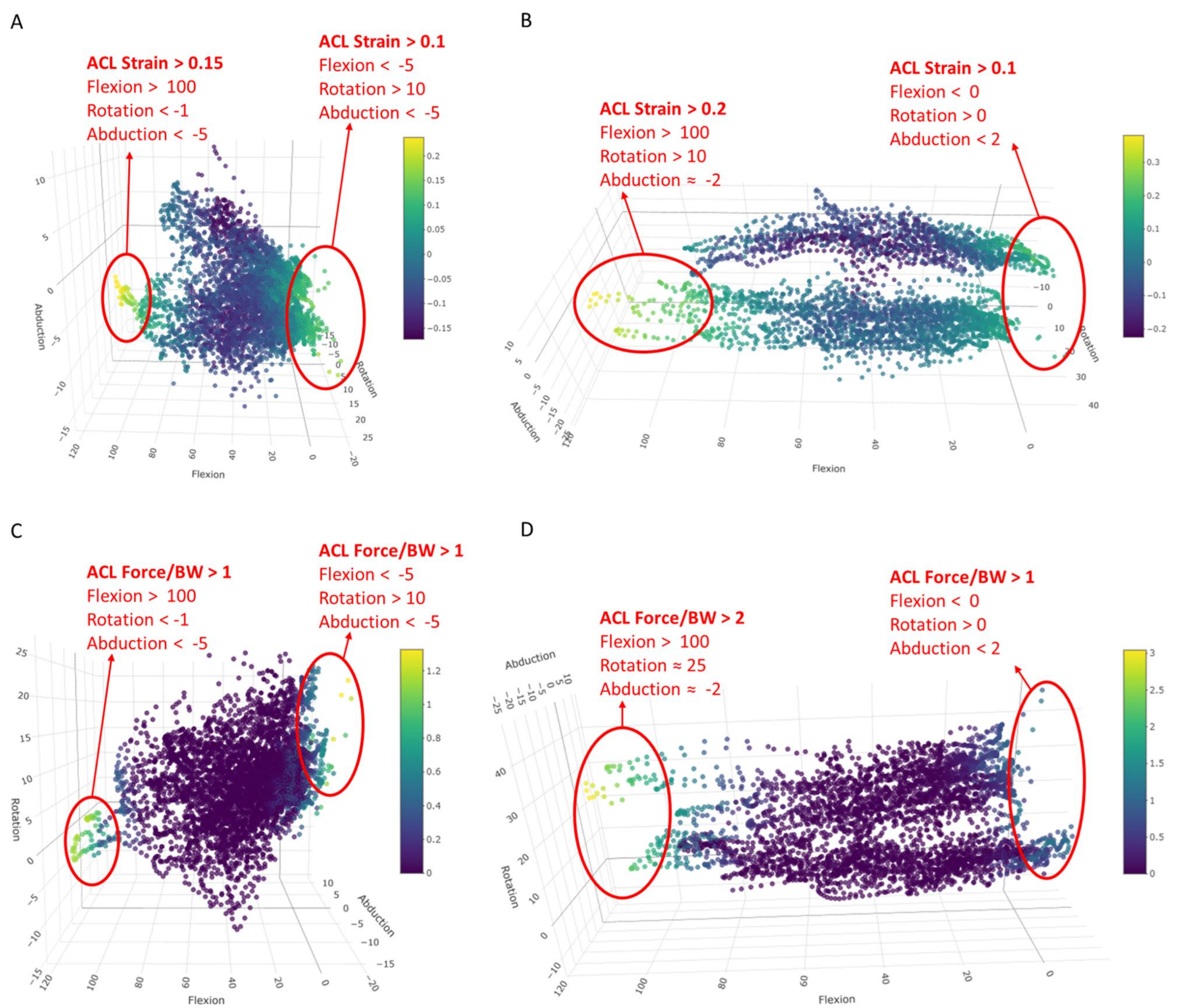

Figure 9. ACL strain and ACL force/BW vs knee angles by sex. A) Men ACL strain, B) women ACL strain, C) men ACL force/BW and D) women ACL force/BW.

analyze all rotations and translations, especially for high-risk activities. Kinematic accuracy may also be affected by skin movement artefacts during motion capture. External knee moments and muscle forces were excluded from the machine learning models, as they were not strong predictors of ACL loading (Daliet et al. 2021), though further investigation could explore additional predictors. Despite excellent R^2 values for cross-validation and test data, a larger dataset (exceeding 12 participants and 9375 observations) would likely improve model accuracy, increase R^2, reduce errors, and capture greater variability in knee biomechanics. While machine learning models offer high accuracy even with non-linear data, they may have limitations when predicting outside their training set, this limitation is minimised in this article by incorporating both daily and high-impact activities, exposing participants to a broad range of motion.

Conclusions

This novel study predicts *in-vivo* ACL length, strain, and force/BW using 42 ML models across 9375 observations per variable. CUBIST, used for the first time in biomechanics, alongside GBM and RF, emerged as the most accurate models (R^2: 0.997–0.992 for cross-validation; 0.984–0.775 for test), effectively estimating ACL variables based on activity, height, weight, gender, and knee flexion, external rotation, and abduction angles, while significantly reducing experimental time and cost. Knee flexion and rotation were the most influential predictors. The study also identifies risky movement patterns associated with high ACL strain and force, aiding in understanding the high incidence of ACL injuries, especially among females. Women showed up to three times higher ACL strain and force/BW than men (3.04 vs 1.33 N/BW), particularly during cross-over cutting, where their knee flexion and rotation increased by ∼20% and 3 times, respectively, compared to males. In contrast, the highest ACL strain and force/BW in men occurred during maximum-effort jumping with knee hyperextension >5°. This novel approach, combining CUBIST models and interactive graphical analysis, enables detection of biomechanical risk patterns, potentially guiding injury prevention in athletes and ACL failure prediction through analysis of joint kinematics. It also has applications in elderly or injured populations for identifying risky knee motions, what can help to improve rehabilitation strategies, and develop customised ACL implants.

Author contributions

CRediT: **Elisa Roldan Ciudad**: Conceptualization; Methodology; Software; Validation; Formal analysis; Investigation; Resources; Data curation; Writing – Original draft; Writing – Review & editing; Visualization; **Neil D. Reeves**: Resources, Writing – review & editing; **Glen Cooper**: Writing – review & editing; **Kirstie Andrews**: Writing – review & editing.

Disclosure statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This research received no external funding.

Data availability statement

All data supporting this article is provided in the manuscript and in the supplementary materials.

References

Agel J, Arendt EA, Bershadsky B. 2005. Anterior cruciate ligament injury in national collegiate athletic association basketball and soccer: a 13-year review. Am J Sports Med. 33(4):524–530. doi: 10.1177/0363546504269937.

Al Kuwaiti A, Nazer K, Al-Reedy A, Al-Shehri S, Al-Muhanna A, Subbarayalu AV, Al Muhanna D, Al-Muhanna FA. 2023. A review of the role of artificial intelligence in healthcare. J Pers Med. 13(6):951. Scopus. doi: 10.3390/jpm13060951.

Andriacchi TP, Dyrby CO. 2005. Interactions between kinematics and loading during walking for the normal and ACL deficient knee. J Biomech. 38(2):293–298. doi: 10.1016/j.jbiomech.2004.02.010.

Baig MM, GholamHosseini H, Moqeem AA, Mirza F, Lindén M. 2017. A systematic review of wearable patient monitoring systems: current challenges and opportunities for clinical adoption. J Med Syst. 41(7):115. doi: 10.1007/s10916-017-0760-1.

Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM. 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophr. 1(1):15030. doi: 10.1038/npjschz.2015.30.

Bien N, Rajpurkar P, Ball RL, Irvin J, Park A, Jones E, Bereket M, Patel BN, Yeom KW, Shpanskaya K, et al. 2018. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. PLoS Med. 15(11):e1002699. Scopus. doi: 10.1371/journal.pmed.1002699.

Blankevoort L, Huiskes R. 1991. Ligament-bone interaction in a three-dimensional model of the knee. J Biomech Eng. 113(3):263–269. doi: 10.1115/1.2894883.

Buller LT, Best MJ, Baraga MG, Kaplan LD. 2015. Trends in anterior cruciate ligament reconstruction in the United States. Orthop J Sports Med. 3(1):2325967114563664. doi: 10.1177/2325967114563664.

Bzdok D, Altman N, Krzywinski M. 2018. Statistics versus machine learning. Nat Methods. 15(4):233–234. doi: 10.1038/nmeth.4642.

Chaaban CR, Berry NT, Armitano-Lago C, Kiefer AW, Mazzoleni MJ, Padua DA. 2021. Combining inertial sensors and machine learning to predict vgrf and knee biomechanics during a double limb jump landing task.

Sensors (Basel). 21(13):4383. Scopus. doi: 10.3390/s21134383.

Chang PD, Wong TT, Rasiej MJ. 2019. Deep learning for detection of complete anterior cruciate ligament tear. J Digit Imaging. 32(6):980–986. Scopus. doi: 10.1007/s10278-019-00193-4.

Daliet OJ, Briem K, Brynjólfsson S, Sigurðsson HB. 2021. Combined effects of external moments and muscle activations on acl loading during numerical simulations of a female model in OpenSim. Appl Sci. 11(24):11971. doi: 10.3390/app112411971.

Dargel J, Gotter M, Mader K, Pennig D, Koebke J, Schmidt-Wiethoff R. 2007. Biomechanics of the anterior cruciate ligament and implications for surgical reconstruction. Strategies Trauma Limb Reconstr. 2(1):1–12. doi: 10.1007/s11751-007-0016-6.

Englander ZA, Baldwin EL, Smith WAR, Garrett WE, Spritzer CE, DeFrate LE. 2019. In vivo anterior cruciate ligament deformation during a single-legged jump measured by magnetic resonance imaging and high-speed biplanar radiography. Am J Sports Med. 47(13):3166–3172. doi: 10.1177/0363546519876074.

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 542(7639):115–118. doi: 10.1038/nature21056.

Fleming BC, Beynnon BD, Renstrom PA, Peura GD, Nichols CE, Johnson RJ. 1998. The strain behavior of the anterior cruciate ligament during bicycling: an in vivo study. Am J Sports Med. 26(1):109–118. doi: 10.1177/03635465980260010301.

Foody JN, Bradley PX, Spritzer CE, Wittstein JR, DeFrate LE, Englander ZA. 2023. Elevated in vivo ACL strain is associated with a straight knee in both the sagittal and the coronal planes. Am J Sports Med. 51(2):422–428. doi: 10.1177/03635465221141876.

Fritz B, Yi PH, Kijowski R, Fritz J. 2023. Radiomics and deep learning for disease detection in musculoskeletal radiology: an overview of novel MRI- and CT-based approaches. Invest Radiol. 58(1):3–13. Scopus. doi: 10.1097/RLI.0000000000000907.

Germann C, Marbach G, Civardi F, Fucentese SF, Fritz J, Sutter R, Pfirrmann CWA, Fritz B. 2020. Deep convolutional neural network-based diagnosis of anterior cruciate ligament tears: performance comparison of homogenous versus heterogeneous knee MRI cohorts with different pulse sequence protocols and 1.5-T and 3-T magnetic field strengths. Invest Radiol. 55(8):499–506. Scopus. doi: 10.1097/RLI.0000000000000664.

Gudigar A, Raghavendra U, Nayak S, Ooi CP, Chan WY, Gangavarapu MR, Dharmik C, Samanth J, Kadri NA, Hasikin K, et al. 2021. Role of artificial intelligence in COVID-19 detection. Sensors (Basel). 21(23):8045. doi: 10.3390/s21238045.

Kalantary S, Jahani A, Jahani R. 2020. MLR and ANN approaches for prediction of synthetic/natural nanofibers diameter in the environmental and medical applications. Sci Rep. 10(1):8117. doi: 10.1038/s41598-020-65121-x.

Kono K, Konda S, Yamazaki T, Tanaka S, Sugamoto K, Tomita T. 2020. In vivo length change of ligaments of normal knees during dynamic high flexion. BMC Musculoskelet Disord. 21(1):552. doi: 10.1186/s12891-020-03560-3.

Leong NL, Petrigliano FA, McAllister DR. 2014. Current tissue engineering strategies in anterior cruciate ligament reconstruction. J Biomed Mater Res A. 102(5):1614–1624. doi: 10.1002/jbm.a.34820.

Lyman S, Koulouvaris P, Sherman S, Do H, Mandl LA, Marx RG. 2009. Epidemiology of anterior cruciate ligament reconstruction: trends, readmissions, and subsequent knee surgery. J Bone Joint Surg Am. 91(10):2321–2328. doi: 10.2106/JBJS.H.00539.

Mak K-K, Pichika MR. 2019. Artificial intelligence in drug development: present status and future prospects. Drug Discov Tod. 24(3):773–780. doi: 10.1016/j.drudis.2018.11.014.

Martin RK, Wastvedt S, Pareek A, Persson A, Visnes H, Fenstad AM, Moatshe G, Wolfson J, Engebretsen L. 2022. Predicting anterior cruciate ligament reconstruction revision: a machine learning analysis utilizing the Norwegian knee ligament register. J Bone Joint Surg Am. 104(2):145–153. Scopus. doi: 10.2106/JBJS.21.00113.

Mesfar W, Shirazi-Adl A. 2006. Knee joint mechanics under quadriceps–hamstrings muscle forces are influenced by tibial restraint. Clin Biomech (Bristol). 21(8):841–848. doi: 10.1016/j.clinbiomech.2006.04.014.

Muqeet M, Malik H, Panhwar S, Khan IU, Hussain F, Asghar Z, Khatri Z, Mahar RB. 2023. Enhanced cellulose nanofiber mechanical stability through ionic crosslinking and interpretation of adsorption data using machine learning. Int J Biol Macromol. 237:124180. doi: 10.1016/j.ijbiomac.2023.124180.

Prodromos CC, Han Y, Rogowski J, Joyce B, Shi K. 2007. A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sport, and a knee injury–reduction regimen. Arthroscopy. 23(12):1320–1325.e6. doi: 10.1016/j.arthro.2007.07.003.

Quatman CE, Hewett TE. 2009. The anterior cruciate ligament injury controversy: is "valgus collapse" a sex-specific mechanism? Br J Sports Med. 43(5):328–335. doi: 10.1136/bjsm.2009.059139.

Rahbar RS, Vadood M. 2015. Predicting the physical properties of drawn Nylon-6 fibers using an artificial-neural-network model. Mater Technol. 49(3):325–332. doi: 10.17222/mit.2013.128.

Roldán E, Reeves ND, Cooper G, Andrews K. 2016. Design consideration for ACL implants based on mechanical loading. Proc CIRP. 49:133–138. doi: 10.1016/j.procir.2015.11.002.

Roldán E, Reeves ND, Cooper G, Andrews K. 2017. In vivo mechanical behaviour of the anterior cruciate ligament: a study of six daily and high impact activities. Gait Posture. 58:201–207. doi: 10.1016/j.gaitpost.2017.07.123.

Roldán E, Reeves ND, Cooper G, Andrews K. 2023a. Can we achieve biomimetic electrospun scaffolds with gelatin alone? Front Bioeng Biotechnol. 11:1160760. doi: 10.3389/fbioe.2023.1160760.

Roldán E, Reeves ND, Cooper G, Andrews K. 2023b. Towards the ideal vascular implant: use of machine learning and statistical approaches to optimise manufacturing parameters. Front Phys. 11. Scopus. doi: 10.3389/fphy.2023.1112218.

Roldán E, Reeves ND, Cooper G, Andrews K. 2024a. 2D and 3D PVA electrospun scaffold evaluation for ligament implant replacement: a mechanical testing, modelling and experimental biomechanics approach. Materialia (Oxf). 33:102042. doi: 10.1016/j.mtla.2024.102042.

Roldán E, Reeves ND, Cooper G, Andrews K. 2024b. Machine learning to mechanically assess 2D and 3D biomimetic electrospun scaffolds for tissue engineering applications: between the predictability and the interpretability. J Mech Behav Biomed Mater. 157:106630. doi: 10.1016/j.jmbbm.2024.106630.

Roldán E, Reeves ND, Cooper G, Andrews K. 2024c. Machine learning to predict morphology, topography and mechanical properties of sustainable gelatin-based electrospun scaffolds. Sci Rep. 14(1):21017. doi: 10.1038/s41598-024-71824-2.

Roldán E, Reeves ND, Cooper G, Andrews K. 2025. Optimising the manufacturing of electrospun nanofibrous structures for textile applications: a machine learning approach. J Textile Inst. 0(0):1–12. doi: 10.1080/00405000.2025.2472089.

Schmitt LC, Paterno MV, Ford KR, Myer GD, Hewett TE. 2015. Strength asymmetry and landing mechanics at return to sport after acl reconstruction. Med Sci Sports Exerc. 47(7):1426–1434. doi: 10.1249/MSS.0000000000000560.

Sharabi M. 2022. Structural mechanisms in soft fibrous tissues: a review. Front Mater. 8 . doi: 10.3389/fmats.2021.793647.

Shimokochi Y, Shultz SJ. 2008. Mechanisms of noncontact anterior cruciate ligament injury. J Athl Train. 43(4):396–408. doi: 10.4085/1062-6050-43.4.396.

Song Y-Y, Lu Y. 2015. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 27(2):130–135. doi: 10.11919/j.issn.1002-0829.215044.

Taylor KA, Cutcliffe HC, Queen RM, Utturkar GM, Spritzer CE, Garrett WE, DeFrate LE. 2013. *In vivo* measurement of ACL length and relative strain during walking. J Biomech. 46(3):478–483. doi: 10.1016/j.jbiomech.2012.10.031.

Tedesco S, Crowe C, Ryan A, Sica M, Scheurer S, Clifford AM, Brown KN, O'Flynn B. 2020. Motion sensors-based machine learning approach for the identification of anterior cruciate ligament gait patterns in on-the-field activities in rugby players. Sensors (Basel). 20(11):3029. Scopus. doi: 10.3390/s20113029.

Utturkar GM, Irribarra LA, Taylor KA, Spritzer CE, Taylor DC, Garrett WE, DeFrate LE. 2013. The effects of a valgus collapse knee position on in vivo ACL elongation. Ann Biomed Eng. 41(1):123–130. doi: 10.1007/s10439-012-0629-x.

Weissler EH, Naumann T, Andersson T, Ranganath R, Elemento O, Luo Y, Freitag DF, Benoit J, Hughes MC, Khan F, et al. 2021. The role of machine learning in clinical research: transforming the future of evidence generation. Trials. 22(1):593. doi: 10.1186/s13063-021-05489-x.

Xiang L, Wang A, Gu Y, Zhao L, Shim V, Fernandez J. 2022. Recent machine learning progress in lower limb running biomechanics with wearable technology: a systematic review. Front Neurorobot. 16:913052. doi: 10.3389/fnbot.2022.913052.

Yang G, Deng J, Pang G, Zhang H, Li J, Deng B, Pang Z, Xu J, Jiang M, Liljeberg P, et al. 2018. An IoT-enabled stroke rehabilitation system based on smart wearable armband and machine learning. IEEE J Transl Eng Health Med. 6:2100510. doi: 10.1109/JTEHM.2018.2822681.

Ye Z, Zhang T, Wu C, Qiao Y, Su W, Chen J, Xie G, Dong S, Xu J, Zhao J. 2022. Predicting the objective and subjective clinical outcomes of anterior cruciate ligament reconstruction: a machine learning analysis of 432 patients. Am J Sports Med. 50(14):3786–3795. Scopus. doi: 10.1177/03635465221129870.

Yoo Y-S, Jeong W-S, Shetty NS, Ingham SJM, Smolinski P, Fu F. 2010. Changes in acl length at different knee flexion angles: an *in vivo* biomechanical study. Knee Surg Sports Traumatol Arthrosc. 18(3):292–297. doi: 10.1007/s00167-009-0932-8.