The Clinical Application of Machine Learning Models for Risk Analysis of Ramp Lesions in Anterior Cruciate Ligament Injuries

Yong-Beom Park,* MD, PhD, Hyojoon Kim,† PhD, Han-Jun Lee,* MD, PhD, Suk-Ho Baek,* MD, Il-Youp Kwak,‡ PhD, and Seong Hwan Kim,*§ MD, PhD

Investigation performed at Chung-Ang University Hospital, Seoul, Republic of Korea

Background: Peripheral tears of the posterior horn of the medial meniscus, known as “ramp lesions,” are commonly found in anterior cruciate ligament (ACL)–deficient knees but are frequently missed on routine evaluation.

Purpose: To predict the presence of ramp lesions in ACL-deficient knees using machine learning methods with associated risk factors.

Study Design: Cohort study (Diagnosis); Level of evidence, 2.

Methods: This study included 362 patients who underwent ACL reconstruction between June 2010 and March 2019. The exclusion criteria were combined fractures and multiple ligament injuries, except for medial collateral ligament injuries. Patients were grouped according to the presence of ramp lesions on arthroscopic surgery. Binary logistic regression was used to analyze risk factors including age, sex, body mass index, time from injury to surgery (≥3 or <3 months), mechanism of injury (contact or noncontact), side-to-side laxity, pivot-shift grade, medial and lateral tibial/meniscal slope, location of bone contusion, mechanical axis angle, and lateral femoral condyle (LFC) ratio. The receiver operating characteristic curve and area under the curve were also evaluated.

Results: Ramp lesions were identified in 112 patients (30.9%). The risk for ramp lesions increased with steeper medial tibial and meniscal slopes, higher knee laxity, and an increased LFC ratio. Comparing the final performance of all models, the random forest model yielded the best performance (area under the curve: 0.944), although there were no significant differences among the models (P>.05). The cut-off values for the presence of ramp lesions on receiver operating characteristic analysis were as follows: medial tibial slope >5.5° (P<.001), medial meniscal slope >5.0° (P<.001), and LFC ratio >71.3% (P=.033).

Conclusion: Steep medial tibial and meniscal slopes, an increased LFC ratio, and higher knee rotatory laxity were observed risk factors for ramp lesions in patients with an ACL injury. The prediction model of this study could be used as a supplementary diagnostic tool for ramp lesions in ACL-injured knees. In general, care should be taken in patients with ramp lesions and its risk factors during ACL reconstruction.

Keywords: ramp lesion; anterior cruciate ligament; meniscal tear; machine learning; random forest

Combined meniscal tears with anterior cruciate ligament (ACL) injuries are common, especially occurring in the posterior horn of the medial meniscus (MMPH).54,55 Among MMPH tears, peripheral meniscocapsular lesions, known as ramp lesions, are difficult to diagnose with magnetic resonance imaging (MRI) or arthroscopic surgery because of their location on the posteromedial aspect of the MMPH, which is difficult to visualize through standard arthroscopic portals.29,54 The pooled diagnostic values of MRI for ramp lesions have been reported as 65.08% for sensitivity and 91.59% for specificity.5,40 Furthermore, the presence of ramp lesions in ACL-injured knees might increase anterior tibial translation and external rotation, as suggested in biomechanical studies.37,56 Thus, the possibility of the misdiagnosis or underdiagnosis of these lesions is concerning,29,54 considering that there are clinical studies that have demonstrated the potential restoration of knee biomechanics after repairing such lesions.37,56 To minimize misdiagnoses or underdiagnoses, identifying the risk factors and understanding the mechanisms of injury associated with ramp lesions are necessary.

Bone morphology and clinical findings around the knee have been considered central risk factors for an ACL injury and failure of ACL reconstruction.11 Several findings, including lateral and medial tibial slopes,19,32,43 meniscal...
slopes,\textsuperscript{32,53} tibial plateau and femoral condyle morphology,\textsuperscript{13,19,44,46,58,61} varus knee alignment,\textsuperscript{5,27,32} bone contusions,\textsuperscript{27,32,63} increased lateral femoral condyle (LFC) ratio (ie, deeper posterior LFC),\textsuperscript{13,31,45,46,61} have been investigated in multiple radiographic or MRI studies as potential risk factors of an ACL injury or ACL reconstruction failure as well as of meniscal injuries. Although these factors have been associated with meniscal tears or ACL injuries in many previous studies, they remain difficult to detect preoperatively.\textsuperscript{5,23,27,37,63} Therefore, a preoperative suspicion of ramp lesions and a careful evaluation during surgery in ACL-injured knees would be vital in obtaining successful outcomes after ACL reconstruction.

Recently, machine learning models have been introduced to predict the risk of certain outcomes or to classify the dataset in accordance with weighing and selecting variables with their calculated importance.\textsuperscript{5,42} In orthopaedic research, the use of machine learning models, including supervised and nonsupervised methods, has been considered promising in predicting survival outcomes,\textsuperscript{12,47} complications,\textsuperscript{11,18} patient assessments,\textsuperscript{14,33} or clinical decision making.\textsuperscript{2,21} Although machine learning models are typically applied to large datasets for the detection of associations among variables, there remain no established criteria for the minimum number of cases and variables for these models. Because of their nature, it is important to evaluate the prediction performance of machine learning models according to data size, which may be collected clinically, more commonly from hospitals than from registry data.\textsuperscript{5,21,22,42} Moreover, in comparing machine learning models and conventional statistical analysis tools, such as logistic regression, the use of a common data size would be clinically valuable.

The random forest model is derived from decision tree analysis, which splits the data into partitions or nodes so that ultimately a previously unseen variable can be accurately assigned to a class.\textsuperscript{1,4,21} The random forest model retains the advantages of decision tree analysis but generates many individual decision trees and ensembles them into a new decision tree after bagging to reduce overfitting of the original classification decision tree, which would be problematic to apply external data.\textsuperscript{1,4,5,21} This random forest model has already been applied for clinical use in medicine to predict outcomes or complications.\textsuperscript{1,17,21}

Deep learning models with neural networks are gaining more popularity than other models.\textsuperscript{10,22,64} Neural networks are known to have advantages in accuracy over statistical models when there is a large number of features in which the relationships are mostly unknown and complex.\textsuperscript{64} Deep learning models with neural networks require massive training and test data samples to utilize their full potential. However, models using neural networks have generally been applied for image analysis, not medical records.\textsuperscript{10,64} In this study, we explored how neural network-based models predict ramp lesions using clinical data.

This study aimed to identify the risk factors for ramp lesions in ACL-injured knees, compare the performance of machine learning models and conventional analysis, and construct an application that can be used in the clinical field. It was hypothesized that machine learning models can predict and improve the diagnostic accuracy of a ramp lesion in ACL-injured knees by combining MRI and multiple variables, including a deeper posterior LFC, varus alignment, and a steeper tibial slope.

METHODS

This retrospective case-control study included 514 cases of primary ACL reconstruction with an arthroscopic evaluation for ramp lesions between June 2010 and March 2019, in accordance with Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines\textsuperscript{3} and quality criteria for artificial intelligence–based prediction in health care.\textsuperscript{7} The inclusion criteria were as follows: primary ACL rupture confirmed by MRI, arthroscopic surgery, and laxity grade >2 on any stress test. A ramp lesion was defined as a longitudinal tear of the meniscocapsular junction or a red zone tear of the MMPH confirmed using arthroscopic surgery.\textsuperscript{29,32,54,55} The exclusion criteria were as follows: (1) fractures around the knee or ACL avulsion fractures (n = 13), (2) previous surgery (n = 11), (3) other combined surgical procedures such as osteotomy (n = 18), (4) multiligament injuries except for medial collateral ligament (MCL) injuries (n = 22), (5) incomplete clinical data (n = 30), and (6) poorly conducted radiographs (n = 58). A total of 362 patients (305 male and 57 female) were enrolled in this study. For this study, we trained in, evaluated, and compared 3 different types of machine learning models for predicting ramp lesions in ACL-injured knees, namely, linear regression, random forest, and neural network. The study was approved by our institutional review board and conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki.

Data Collection

Patient characteristics including age, sex, body mass index (BMI), mechanism of injury, and time between surgery and injury (ie, <3 or ≥3 months) were obtained. MRI
examinations were performed on a 3.0-T machine (Magnetom Aera; Siemens) using an 8-channel knee coil with a 1.5-mm slice thickness. A ramp lesion on MRI was confirmed by an experienced musculoskeletal radiologist and an orthopaedic surgeon specializing in knees who were not present during surgery. It was defined as a high signal change or separation in the meniscocapsular junction of the MMPH on sagittal images.23,32 Ramp lesions were definitively confirmed during arthroscopic surgery 29 (Figure 1).

Bone contusions were evaluated on MRI and documented as being located on the LFC, lateral tibial plateau, medial femoral condyle, or medial tibial plateau (MTP).32 The medial and lateral tibial and meniscal slopes on MRI were measured using previously reported methods.26,32,53 A deep sulcus sign, Segond fractures, and MCL injuries, regardless of the grade, were also assessed on MRI and plain radiography.48,52 The LFC ratio was measured on plain radiography using the method described by Pfeiffer et al45,46 (Figure 2).41 The long axis of the femoral shaft (line 1) was determined by a line through the center of 2 circles on the femoral shaft. The axis of the femoral condyle (line 2) was determined by a line between the most posterior and most anterior points of the lateral condyle. The distance from the intersection of these lines to the posterior end of the condyle was divided by the total anteroposterior length of the condyle. The angle between lines 1 and 2 was defined as the condyle flexion angle.

Surgical Procedure

All surgical procedures were performed at a single institution. Systematic arthroscopic evaluations were performed...
with 30° and 70° arthroscopes during ACL reconstruction. After a standard arthroscopic evaluation through the anterolateral and anteromedial portals, exploration of the posteroomedial compartment through the intercondylar space using a 30° arthroscope was performed. A 70° arthroscope was then used in situ to evaluate the MMPH and posteroomedial corner with thorough probing, confirming the ramp lesion. After final exploration, the ramp lesions were trephined using a motorized shaver and repaired using a suture hook through the posteroomedial portal with absorbable sutures (PDS; Ethicon). Repair of the ramp lesions was not performed if they were <5 mm in length and were stable on probing.

Statistical Analysis

Data were analyzed using SPSS (version 19.0; IBM), R (version 4.0.3; R Foundation), Scikit-learn and TensorFlow in Python modules, and G*Power (version 3.1.5). The R packages that we used were caret, pROC, ROCR, randomForest, caTools, devtools, repretree, and e1071. To compare the mean values, data were analyzed using the Mann-Whitney U test, independent t test, paired t test, chi-square test, or Wilcoxon signed-rank test after the Shapiro-Wilk test for normality of the distribution. The predictive values were compared using logistic regression analysis with machine learning models, namely, random forest and neural network. The predictive factors were age, BMI, sex, time from injury to surgery (>3 or <3 months), bone contusions, presence of Segond fractures, presence of MCL injuries, pivot-shift grade (high vs low), LFC ratio, injury mechanism (contact vs noncontact), presence of varus knee alignment (>6° or ≤6°), amount of anterior translation on stress radiography, and medial/lateral tibial and meniscal slopes on MRI. The receiver operating characteristic (ROC) curve and area under the curve (AUC) were plotted to identify the cut-off radiographic measurements for ramp lesions with the maximum point of the Youden index, which indicated the farthest point from the reference line in ROC curves. The sensitivity, specificity, positive predictive value, and negative predictive value were also calculated using ROC analysis. Furthermore, ROC curve comparisons were performed between each model set using the DeLong test for 2 correlated ROC curves.

Using Scikit-learn and TensorFlow in Python modules, the complete dataset was randomly divided into 2 subsets as follows: a training sample (254 cases [70.2%]) and a test sample (108 cases [29.8%]). This was done to estimate the predictive ability of each model. For each model, we performed k-fold cross-validation with k = 100 and selected a model with 90th percentile accuracy among 20 runs. For conventional statistical analysis, binary stepwise logistic regression with backward elimination was performed to evaluate the risk factors for ramp lesions using the Akaike information criterion.

For random forest analysis, multiple trees were grown to create a final random forest model. Each individual tree grew from a bootstrap sample (ie, a random sample selected with replacement [“bagging”]) and was an unpruned decision tree grown using the Gini impurity measure. The grown individual trees were grouped as an ensemble, creating a final tree model, which was defined as a random forest.

For neural network analysis, we used a total of 3 dense layers. The first 2 layers used the rectified linear unit activation function, which is one of the most successful and widely used functions in neural networks. The last layer used the sigmoid activation function, which was relevant for making a binary decision. The neural network started with 20 neurons, similar to the number of variables in our dataset, and the last layer ended with a single neuron to make a binary decision. Increasing the number of layers beyond 3 did not improve the model’s accuracy. Contrarily, it started to degrade because of possible overfitting. Having too many layers resulted in decreasing back-propagating errors, which made the learning process ineffective. For the loss function used to minimize the prediction error, we used the binary cross-entropy loss function, which was suitable for a neural network model with binary output.

ROC analysis on a test set was used to evaluate a model’s accuracy and the importance of included variables and factors. We plotted ROC curves and calculated the AUC to further analyze and compare the diagnostic accuracy of the different models. The DeLong test was used to enable an ROC curve comparison between 2 different models. To determine the most important variables and factors that influence the random forest model, we estimated coefficients by calculating and creating partial dependence plots. Then, the final prediction model was established to improve the diagnostic value by combining the best machine learning model and MRI outcomes.

An a priori sample size calculation was performed using the goodness-of-fit test with an alpha error of 0.05 and power of 0.9. We assumed 0.7 for the diagnostic accuracy of ramp lesions on MRI according to previous studies and 0.8 for the diagnostic accuracy of ramp lesions with a machine learning model. The required sample size was calculated as 221 with an effect size of 0.2182179.

Furthermore, we calculated the sample size and the number of ramp lesions/cases required to target performance measures (AUC and observed/expected [O/E] calibration score) in an external validation study of a ramp lesion detection model using the closed-form formula proposed by Riley et al. The proportion of events was assumed to be 0.309, as suggested from our experimental data. The target AUC was set at 0.8 with a 95% CI of 0.1 (related standard error of AUC = 0.0255), and the target O/E calibration score was set at 1.0 with a 95% CI of 0.2 (related standard error of O/E score = 0.0510).

Statistical significance was set at P < .05. All radiographs or MRI scans were reviewed twice by 2 musculoskeletal radiologists or an orthopaedic surgeon specializing in the knee with a 2-week interval. The interobserver and intraobserver reliabilities of measurements were assessed using the kappa value for agreement or the intraclass correlation coefficient for consistency.
RESULTS

Patient characteristics, according to the presence of ramp lesions, are summarized in Table 1. Among 362 patients who underwent ACL reconstruction, a ramp lesion was identified in 112 patients (30.9%) during arthroscopic surgery. However, ramp lesions were identified in 128 patients (35.4%) on MRI. The diagnostic value of MRI for the presence of ramp lesions had an accuracy of 82.3%, sensitivity of 78.6%, specificity of 84.0%, positive predictive value of 68.8%, and negative predictive value of 89.7%. Among 112 patients with ramp lesions, 106 cases (94.6%) were repaired, and 6 cases (5.4%) were left in situ.

Associated Risk Factors for Ramp Lesions on Chi-Square Analysis

On chi-square analysis, the associated risk factors for a ramp lesion in ACL-deficient knees were chronic ACL injuries (odds ratio [OR], 1.857 [95% CI, 1.055-3.270]; P = 0.030), a noncontact injury mechanism (OR, 1.792 [95% CI, 1.099-2.922]; P = 0.019), a high-grade pivot shift (OR, 4.444 [95% CI, 2.738-7.215]; P = < 0.001), a deep sulcus sign (OR, 2.011 [95% CI, 1.193-3.391]; P = 0.008), medial femoral condyle bone contusions (OR, 1.633 [95% CI, 1.000-2.669]; P = 0.049), and MTP bone contusions (OR, 3.171 [95% CI, 1.997-5.035]; P = 0.000) (Table 1). There were also significant differences in the LFC ratio and medial tibial and medial meniscal slopes between groups.

Results of Logistic Regression Analysis

Factors associated with ramp lesions were identified using logistic regression analysis with backward elimination (Table 2). The risk for ramp lesions in ACL-deficient knees increased with chronic injuries (≥3 months), a steep medial tibial slope, a gradual lateral tibial slope, a high-grade pivot shift, MTP bone contusions, a deep sulcus sign, and an increased LFC ratio, indicating a deeper LFC directed posteriorly. Among them, an increased LFC ratio was the most significant factor based on the OR, followed by a high-grade pivot shift (Table 2).

Comparison of Diagnostic Values Between Machine Learning Models and Logistic Regression Model

The results of the random forest, neural network, and logistic regression models are shown in Table 3. To identify the best hyperparameters for the random forest model, we performed a grid search and subsequently established a model with the following hyperparameters: 200 trees and a minimum of 2 variables at each split, with a minimum terminal size of 1 (Table 3). The prediction error rate for the test set using the random forest model was 6.3%. In the importance plot for the random forest model, pivot-shift grade, side-to-side laxity, medial meniscal and tibial slopes, and MTP bone contusion were the important factors (Figure 3), followed by BMI and LFC ratio. The LFC ratio was not the highest influencing factor in the random forest model, although it yielded the highest OR in the logistic regression model. MTP bone contusion and pivot-shift grade were the important factors for both the random forest and logistic regression models. In our partial dependence plots, an LFC ratio of approximately 0.71 to 0.75 (Appendix Figure A1A, available in the online version of
this article), a medial tibial slope of 7° to 9° (Appendix Figure A1B), and a medial meniscal slope of 5° to 6° (Appendix Figure A1C) were the cut-off values in the random forest model to predict ramp lesions.

Various combinations of layers and nodes were tested to determine the best fit in the neural network model. Ultimately, our neural network was constructed with 20 nodes, 5 nodes, and 1 node in the first, second, and third layers, respectively. The rectified linear unit activation function was used for the first 2 layers, and the sigmoid activation function was used for the last layer to produce a single classification value. The prediction error rate for the test set using the neural network model was 7.0% (Appendix Figure A2, available in the online version of this article).

To compare the final performance of all models, we calculated the accuracy, sensitivity, specificity, and AUC (Table 3 and Figure 4). The random forest model yielded the best performance in terms of accuracy, sensitivity, and prediction error, but there were no significant differences among the models \( (P > .05) \). Although the logistic regression model showed the highest AUC over the neural network and random forest models, the accuracy and prediction error for the former were worse than those for the random forest model.

The kappa value of interobserver and intraobserver reliability for ramp lesions was 0.806 \( (P < .001) \) and 0.854 \( (P < .001) \), respectively, indicating good agreement between the raters. The intraclass correlation coefficient values for the reliability of radiographic measurements ranged from 0.75 to 0.84, indicating good agreement among raters.

ROC Curve Analysis for Presence of Ramp Lesions

The AUC, sensitivity, and specificity were calculated to estimate the diagnostic accuracy of each radiographic measurement for ramp lesions, which were clinically significant among the analyses (Table 4 and Figure 5). When comparing among the factors, the ROC curve for side-to-side laxity was significantly different from that for the medial tibial and meniscal slopes \( (P < .001) \). Although the medial meniscal slope was not a significant factor in the logistic regression model, it was significant on ROC analysis \( (\text{AUC}: 0.644; P < .001) \). In this analysis, a medial tibial slope \( >5.5° \) (sensitivity: 81.9% [95% CI, 73.2%-89.3%]; specificity: 43.1% [95% CI, 37.3%-49.3%]), an LFC ratio \( >71.3% \) (sensitivity: 24.1% [95% CI, 16.5%-33.1%]; specificity: 89.6% [95% CI, 85.1%-93.3%]), and a medial meniscal slope \( >5.0° \) (sensitivity: 62.5% [95% CI, 52.9%-71.5%]; specificity: 66.0% [95% CI, 59.8%-71.9%]) were cut-off values for the presence of ramp lesions.

Final Prediction Model Including MRI Findings and Random Forest Model

To improve the diagnostic values of MRI, the final prediction model was established with a combination of each models and MRI findings (Table 5). The random forest model had the highest values, with an accuracy of 91.7%, sensitivity of 85.2%, and specificity of 93.9% (Figure 6).
The STARD flow diagram for the test set using the final random forest model is shown in Figure 7 (http://rampacl.com).

In the sample size calculation for external validation, we required a sample size of 353 (and 109 events) by assuming that the proportion of events was 0.309 and the AUC was 0.8 with a targeted 95% CI of 0.1. For the O/E calibration score, 897 participants (and 269 events) were required to aim for a 95% CI of 0.2 according to the study of Riley et al49 (Table 6).

DISCUSSION

This study notably found that the risk of ramp lesions was higher in ACL-injured knees with steeper medial tibial and meniscal slopes, higher knee laxity including pivot-shift grade and side-to-side laxity, body mass index (BMI), and lateral femoral condyle (LFC) ratio were the important factors.

An increased LFC ratio has been reported to be associated with ACL injuries and ACL reconstruction failure.16,19,45,46 This anatomic feature could alter the gait and loading mechanics of the tibiofemoral joint and increase the length and anisometry in the anterolateral ligament complex, which might increase anterolateral rotatory laxity.25,45,46 As suggested by Pfeiffer et al,45,46 an increased LFC ratio, which may be more anisometric and oval shaped, might lead to increased ligamentous laxity in the lateral or anterolateral ligament complex. Moreover, they reported that a decreased contact area due to an oval-shaped femur might also increase rotatory laxity.46 Similar results were also found in the study by Gaillard.
et al.\textsuperscript{16} who reported that a greater anteroposterior length of the femoral condyle than that of the tibial plateau is associated with an increased risk of meniscal tears in knees with an acute ACL injury. However, LFC anatomy was also suggested as a secondary risk factor for ACL failure because of no significant differences in the LFC ratio between primary and ancillary ACL reconstruction in a study by Grassi et al.\textsuperscript{20} although it was significant in the multiple ACL failure group. Although the association between ramp lesions in ACL injuries and an increased LFC ratio has been discussed recently,\textsuperscript{31} the lack of literature hampers the establishment of an association between LFC ratio and ramp lesions. In our study, an increased LFC ratio was found to be a significant risk factor for ramp lesions (OR, 378.418 [95% CI, 1.356-105564.8]; \(P = .035\)), and the cut-off value was found to be \(71.3\%\) (AUC: 0.571; sensitivity: 24.1%; specificity: 89.6%; \(P < .033\)), although its effect on the machine learning model was limited (Figure 2A). The cut-off value for the presence of ramp lesions was slightly higher than the cut-off value that increased the risk of ACL injuries (ie, LFC ratio of 68%) in a previous study by Pfeiffer et al.\textsuperscript{45} Because a concomitant injury in the meniscus, particularly ramp lesions, could be considered a more severe knee injury than an isolated ACL injury, the cut-off value for the LFC ratio in this study might be higher than that for the ACL injury risk, along with increased rotatory laxity caused by the deeper posterior LFC. Additionally, a high-grade pivot shift (grade \(\geq 2\)) and side-to-side laxity were found to be significant factors for the presence of ramp lesions in the logistic regression and random forest models (Figure 3A); thus, higher laxity due to the deeper posterior LFC could increase the risk for ramp lesions in ACL-injured knees.

Anterior laxity after an ACL injury could induce engagement of the medial meniscus in the femoral condyle, resulting in high-stress loading on the meniscocapsular junction.\textsuperscript{36,38,53} An increased tibial slope could induce an increase in anterior translation of the tibia in ACL-deficient knees,\textsuperscript{26,28,58} which could in turn increase the incidence of ramp lesions.\textsuperscript{26,32,38} In a study by Okazaki et al.,\textsuperscript{43} an increase in the medial tibial slope and a shallow concave shape of the MTP were identified as possible risk factors for medial meniscus root tears. This is because of decreased resistance in anterior translation of the medial tibia, resulting in greater loading stress on the

\begin{table}[h]
\centering
\caption{Comparison of Models After Inclusion of MRI Findings\textsuperscript{a}}
\begin{tabular}{|l|l|l|l|l|l|}
\hline
 & AUC (95\% CI) & Sensitivity, \% & Specificity, \% & Accuracy (95\% CI) & Prediction Error, \% \\
\hline
Logistic regression & 0.927 (0.873-0.981) & 84.8 & 89.5 & 0.881 (0.836-0.925) & 13.3 \\
Neural network & 0.906 (0.844-0.968) & 75.7 & 91.7 & 0.862 (0.798-0.927) & 12.6 \\
Random forest & 0.944 (0.877-1.000) & 85.2 & 93.9 & 0.917 (0.873-0.954) & 2.6 \\
\hline
\end{tabular}
\textsuperscript{aAUC, area under the curve; MRI, magnetic resonance imaging.}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Comparison of receiver operating characteristic (ROC) curves for prediction by individual covariates. When comparing ROC curves among the factors, the ROC curve for side-to-side laxity was significantly different from that for the medial tibial and meniscal slopes (\(P < .001\)).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Comparison of receiver operating characteristic (ROC) curves for the prediction performance of the models including magnetic resonance imaging findings. The random forest model yielded the largest area under the curve (AUC).}
\end{figure}
posteromedial compartment of the knee. A steeper medial tibial slope might also induce excessive anterior sliding of the MTP after an initial pivot-shift injury as part of the contrecoup mechanism.27,32,63 In this study, as demonstrated by a bump in the machine learning model, an increased medial meniscal slope was found as a risk factor for ramp lesions, consistent with the study findings of Song et al.53 The results of this study were also consistent with those of previous studies.31,32,38,43 A medial tibial slope of 5.5° was found to be a significant risk factor for ramp lesions in this study. Briefly, accompanying an increased LFC ratio, increased anterior and rotatory laxity occurred initially, confirmed by the greater side-to-side laxity and high-grade pivot shift in this study. These were followed by a secondary injury (during the contrecoup mechanism) with excessive anterior sliding of the MTP because of a steeper medial tibial slope.27,32,63 Furthermore, the total range of the rotational arc might be increased during a knee injury. This could increase loading on the MMPH and menisocapsular junction and affect the medial meniscal slope, which may result in a ramp lesion in the ACL-injured knee. It can be assumed that a ramp lesion in an ACL-injured knee possibly represents a more severe knee injury than an isolated ACL injury; thus, combined lateral procedures to secure rotatory laxity, such as lateral tenodesis or anterolateral ligament reconstruction, might be necessary.27,32,63

In this study, patients with a chronic ACL injury (≥3 months from injury) were found to have a higher incidence of ramp lesions (42.6% [26/61]) than those with an acute ACL injury (28.6% [86/301]) (P = .030) (Table 1). Many previous studies have also reported that ramp lesions might increase over time.9,36,55 Further, DePhillipo et al9 reported no difference in clinical outcomes, knee stability, and return-to-sports rates between the ACL reconstruction with ramp lesion repair group and isolated ACL reconstruction group, despite the increased incidence of ramp lesions in patients with a chronic ACL injury. Thus, ramp lesions should be carefully assessed and repaired, especially in chronic ACL injuries, as they are difficult to treat through standard anterior arthroscopic surgery.29,54 Systematic arthroscopic visualization of the posteromedial knee compartment, as performed in this study, would be helpful in diagnosing ramp lesions29,32,54,55 because the incidence of ramp lesions in this study (30.9%) is relatively higher than that in previous studies.23,36,54,55 However, although chronicity was significant on chi-square and logistic regression analyses, its importance in the machine learning model was found to be limited (Figure 3A).

In our study, the machine learning models, including the random forest and neural network models, showed similar

---

### Table 6

Sample Size Calculation for External Validation

<table>
<thead>
<tr>
<th>Assumed Value</th>
<th>Targeted 95% CI</th>
<th>No. of Participants/Events Required to Achieve 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-statistic (AUC)</td>
<td>0.8</td>
<td>0.1 (0.0255)</td>
</tr>
<tr>
<td>O/E calibration score</td>
<td>1.0</td>
<td>0.2 (0.0510)</td>
</tr>
</tbody>
</table>

*The codes for the complete model-building pipeline are available at https://github.com/Medical-ML/ramp-acl. AUC, area under the curve; O/E, observed/expected.
performance to that of the conventional logistic regression model in terms of the AUC, but lower prediction errors were found in the random forest and neural network models than in the logistic regression model. The conventional logistic regression model is prone to overfitting of training data when used as a prediction model, often resulting in poorer performance when presented with new data (ie, test data), which makes it difficult to use clinically. Neural network models have advantages in predicting outcomes by utilizing complex relationships between features that are difficult to identify or understand even by field experts. The random forest model is an algorithm that generally improves on the disadvantages of decision tree analysis in which numerous classification trees are developed, each using a random subset of the available training data, and all trees form an ensemble. This randomness often circumvents the overfitting observed in traditional models. Overfitting means that the model fits too closely to the training dataset, making the model less generalizable. This results in the model performing poorly against different datasets; thus, it is generally known to be more appropriate for external data analyses. The random forest model showed the lowest prediction error rate when applied to the test set. The sample size of this study might be too small to construct a neural network model because of the need for several hidden layers, nodes, or weights for each node. In summary, the random forest model was selected to establish the final prediction model to predict the presence of ramp lesions, which has potential clinical applications (http://rampacl.com). Thus, the combination of direct findings on MRI and indirect findings in the final prediction model showed higher diagnostic accuracy than that of MRI alone as well as that of previous studies.

This study has several limitations. The first is its retrospective nature. Second, other factors that might be associated with ramp lesions, such as anterolateral ligament injuries, were not included in this study. A combined anterolateral ligament injury has been reported as a risk factor for ramp lesions in previous studies. Third, the proposed correlation of femoral morphology and ramp lesions is an assumption. Although a deeper posterior LFC was found to be a significant factor for ramp lesions in this study, other femoral morphometric findings may also be risk factors for ramp lesions if they were reported as risk factors for an ACL injury. A narrow LFC has also been reported as a risk factor for ACL injuries, as has a short flat surface of the LFC with small anteroposterior distances of the tibial plateaus. The increased risk of ACL injuries associated with the aforementioned factors might similarly predispose affected patients to ramp lesions. Fourth, the performance of machine learning models was found to be similar to that of conventional statistical analysis, although the versatile application of machine learning models was an advantage. However, there might not be an adequate sample size to obtain excellent performance for machine learning models. Further studies are necessary for the improvement of machine learning model performance, either involving a larger sample size or a specific optimization method. This is particularly true for the neural network model, which typically requires a greater sample size than in this study to maximize its use. Furthermore, a wide 95% CI and a high OR for the LFC ratio on logistic regression analysis were found because of numbers with decimal points; care should be taken to interpret this result. If one uses the value as a percentage rather than a number with a decimal point, the OR and its 95% CI would be smaller. Fifth, although the final prediction model showed higher diagnostic values than MRI, the sensitivity was still lower than 90%. Because the accuracy for this model was still limited in its use for diagnostic confirmation, a larger dataset with an evaluation of other risk factors would be needed to improve this machine learning model. Furthermore, varus alignment was dichotomized in this machine learning model, even though it was measured as a continuous variable. If varus alignment could be used as a continuous value with a larger dataset, it might be significant, which can improve the performance of the machine learning model. Furthermore, external validation using this machine learning model would be better to assess exact performance. This study showed a potential to establish a machine learning model using small data, representing “real-world” data. Further studies would be needed in terms of external validation, a larger dataset to improve model performance, and automated measurements of radiographic parameters by combining the various models (eg, convolutional neural network for image analysis and then combined random forest with clinical data). There might be a potential to improve this model.

CONCLUSION

Steep medial tibial and meniscal slopes, an increased LFC ratio, and higher knee rotational laxity were found to be risk factors for ramp lesions in patients with an ACL injury. The prediction model of this study could be used as a supplementary diagnostic tool for ramp lesions in ACL-injured knees. Care should be taken in patients with ramp lesions and risk factors during ACL reconstruction.

ORCID iDs

Yong-Beom Park https://orcid.org/0000-0002-3741-2311
Seong Hwan Kim https://orcid.org/0000-0001-5471-5747

REFERENCES


For reprints and permission queries, please visit SAGE’s Web site at http://www.sagepub.com/journals-permissions