Re-tear after arthroscopic rotator cuff tear surgery: risk analysis using machine learning

Shinohara, Issei ; Mifune, Yutaka ; Inui, Atsuyuki ; Nishimoto, Hanako ; Yoshikawa, Tomoya ; Kato, Tatsuo ; Furukawa, Takahiro ; Tanaka, Shuy…

(Citation)
Journal of Shoulder and Elbow Surgery,33(4):815-822

(Issue Date)
2024-03-12

(Resource Type)
journal article

(Version)
Accepted Manuscript

Re-tear after arthroscopic rotator cuff tear surgery: risk analysis using machine learning

Abstract

Background: Postoperative rotator cuff re-tear after arthroscopic rotator cuff repair (ARCR) is still a major problem. Various risk factors such as age, gender, and tear size have been reported. Recently, magnetic resonance imaging (MRI)-based stump classification was reported as an index of rotator cuff fragility. Although stump type 3 is reported to have a high re-tear rate, there are few reports on the risk of postoperative re-tear based on this classification. Machine learning (ML), an artificial intelligence technique, allows for more flexible predictive models than conventional statistical methods and has been applied to predict clinical outcomes. In this study, we used ML to predict postoperative re-tear risk after ARCR.

Methods: The retrospective case-control study included 353 patients who underwent surgical treatment for complete rotator cuff tear using the suture-bridge technique. Patients who initially presented with re-tears and traumatic tears were excluded. In study participants, after the initial tear repair, rotator cuff re-tears were diagnosed by MRI; Sugaya classification types IV and V were defined as re-tears. Age, gender, stump classification, tear size, Goutallier classification, presence of diabetes, and hyperlipidemia were used for ML parameters to predict the risk of re-tear. Using Python's Scikit-learn as an ML library, five different AI models (logistic regression, random forest, AdaBoost, CatBoost, LightGBM) were trained on the existing data, and the prediction models were applied to the test dataset. The performance of these ML models was measured by the area under the receiver operating characteristic curve (AUC). Additionally, key features affecting re-tear were evaluated.

Results: The AUC for logistic regression was 0.78, random forest 0.82, AdaBoost 0.78, CatBoost 0.83, and LightGBM 0.87, respectively for each model. LightGBM showed the highest score. The important factors for model prediction were age, stump classification, and tear size.

Conclusions: The ML classifier model predicted re-tears after ARCR with high accuracy, and the AI model showed that the most important characteristics affecting re-tears were age and imaging findings, including stump classification. This model may be able to predict postoperative rotator cuff re-tears based on clinical features.

Study design: Prognosis Study (Case-control study).

Level of evidence: III

Keywords: arthroscopic rotator cuff repair; artificial intelligence; feature importance; LightGBM; machine learning; re-tear; SHAP; stump classification

Introduction

Postoperative re-tear is still a problem in arthroscopic rotator cuff repair (ARCR) for degenerative rotator cuff tears (RCTs). The reported re-tear rate after ARCR varies depending on the suture method, ranging from 5-92%.[6,5,9,13,14,29,39] Assessment of risk factors is important since re-tears significantly reduce postoperative function and require reoperation.[39] Risk factors for postoperative re-tears include age,[2,3,11] tear size,[3,4,7,11] fatty degeneration,[15,16,24] and suturing technique.[5] Recently, stump classification using the coronal view of T2 fat suppression on magnetic resonance imaging (MRI) was proposed as an indicator of rotator cuff fragility.[23] Comparing the signal intensity of the deltoid (D) and the rotator cuff tear (C), C<D is classified as type 1, C=D as type 2, and C>D as type 3.[23] Stump type 3 was reported to have a significantly higher postoperative re-tear rate after ARCR, suggesting that stump classification may be an indicator of rotator cuff fragility.[39] It has also been suggested that advanced glycation end-products (AGEs), which rise with aging and diabetes mellitus (DM), are associated with tendon fragility.[37] Inflammation and degeneration caused by oxidative stress and abnormal collagen cross-linking due to the accumulation of AGEs affect stump classification by MRI images. There are few reports taking stump classification into account that may be useful for predicting re-tears after ARCR. In this study, we focused on the analysis of clinical data by machine learning (ML), which has recently attracted attention in the field of orthopaedics.[21] ML, an artificial intelligence (AI) technique, is a method capable of incorporating patient-related variables into predictive models and providing individualized risk assessments.[32] It allows for more flexible predictive models than conventional statistical methods and has been applied to predict clinical outcomes.[32] ML has been applied to a variety of fields: sports medicine,[25,26] joint surgery,[21,32] and spine surgery,[33] and has been reported as an algorithm for predicting factors affecting clinical outcomes and improvements. There are also reports on using ML to predict RCTs in terms of assessing important clinical features[27] and predicting costs.[18] On the other hand, there are no reports on the inclusion of stump classification in ML models to predict re-tears after ARCR.

The purposes of this study are twofold; first, to evaluate the predictive accuracy for re-tears after ARCR by applying ML to clinical data, and second, to evaluate the features that the AI determines to be important in predicting re-tears, including stump classification. This study was based on the hypothesis that a classifier generated by ML would predict postoperative re-tears after ARCR with high accuracy and stump classification may be an important feature in predicting re-tear.

Materials and Methods

Ethical approval

This study was approved by the appropriate review board, and informed consent was obtained from all patients involved.

Data collection

Patients who underwent ARCR for degenerative complete rotator cuff tears from April 2017 to June 2021 at our institution or affiliated institutions were included. ARCR was performed by two surgeons, Y.M. and M.M., using the suture bridge technique. Reoperations, trauma, and patients who required patch augmentation for rotator cuff repair were excluded from this study. Traumatic tears were defined to include trauma to the symptomatic shoulder, such as falls, impacts, and sudden extensions.[34] MRI was used to identify study participants who suffered rotator cuff re-tears, with Sugaya classification types IV and V defined as re-tears.[38] The parameters for ML were age, gender, medical history (DM, hyperlipidemia), stump classification (Fig. 1),[23] tear size,[10] and fatty degeneration (Goutallier classification).[17]

Statistical analysis

Each patient parameter is expressed as mean ± standard deviation. To compare patient backgrounds with and without re-tears, the Mann–Whitney U test was used to compare two variables (e.g., gender) and Fisher's exact test to compare multiple variables (e.g., tear size). Statistical significance was set at $p < 0.05$.

Machine learning

The data collection and ML workflow are shown in Fig. 2. Five supervised algorithms were applied to validate clinical data[20,22] (logistic regression, random forest,[1] adaptive boost (AdaBoost),[35] CatBoost,[20] and light gradient-boosting machine (LightGBM), which is a modified gradient boosting decision tree,[42]) were used as ML algorithms to predict rotator cuff re-tears after ARCR, and the predictions were compared. The logistic regression model is a widely used multivariate analysis approach in medical research. The remaining models are general ensemble methods that combine multiple simple tree models and have been proven to make reliable predictions.[30] Random forest is a method that uses ensemble decision trees to extract random subsets from the data with replacement, allowing all data to be used for training and validation while avoiding the tendency of decision trees to overfit models.[1] In brief, it is a method that attempts to obtain better predictions by using multiple training models and performing majority voting on the results.[1] On the other hand, Adaboost, Catboost, and LightGBM are gradient boosting methods, which take over the errors from the previous decision tree calculation and correct them. AdaBoost is a learning algorithm that feeds back errors made in training and iteratively learns to improve accuracy.[35] Feedback reduces the error of the ML and allows a better accuracy rate to be reached. The approach has been applied to the data analysis of COVID-19.[35] CatBoost is a ML algorithm that can highly process categorical variables and is widely used for big data analysis.[20] LightGBM is a model that greatly improves the computation time due to scanning all the sample points of each feature when finding the optimal split point in the boosting process.[42] LightGBM increases computational speed by growing the decision trees used, reducing memory footprint, improving classification accuracy, and efficiently preventing overfitting.[42] Scikit-learn, a free ML library for Python,[31] was used to implement these supervised algorithms. Patient data were randomly divided into training samples (70%) used for hyperparameter tuning to generate ML models, and validation samples (30%) to verify the performance of each model. After the optimal hyperparameters for each ML algorithm were determined in the training sample data, the prediction accuracy (percentage of correct answers for all data) of re-tear in each model for the test data was evaluated. For each ML model, the accuracy and the area under the curve (AUC) obtained from the receiver operating characteristic (ROC) were calculated. AUC in ML indicates accuracy of the classifier. For each endpoint, 95%

confidence intervals (CI) were calculated using the bootstrap method.[30] The bootstrap method is an iterative resampling method used to estimate key statistics, such as the mean and standard deviation, by resampling and resubstituting the data set.[30] In addition, key values of each prediction parameter were computed using two different algorithms to visualize the basis for the ML model's decisions: permutation feature importance is defined as the amount by which the model score decreases when one feature value is randomly shuffled;[12] the Shapley additive explanation (SHAP) value is defined as the contribution of each feature to the model prediction based on game theory.[40] Briefly, it is a method for determining the contribution of each variable (feature) to the predicted results of the ML model.[40]

Results

Study participants and statistical analysis

Of the 582 cases who underwent ARCR at our institution or affiliated institutions, 353 were finally included after excluding re-tears (12 cases), traumatic tears (182 cases), and patients who required patch augmentation. In the study participants, re-tears were observed in 45 cases (12.7%); the mean time to postoperative re-tear was 9.4 ± 3.7 months. A statistical analysis of patient background based on the presence or absence of rotator cuff re-tears is shown in Table 1.

Prediction of rotator cuff re-tear in each ML model

Fig. 3 shows a heat map representing the correlation between each parameter and rotator cuff re-tear. Warm colors indicate a positive correlation, while cold colors indicate a negative correlation. The heat map showed that DM, stump type, tear size, and fatty degeneration of the rotator cuff were positively correlated with re-tear. The accuracy and AUC for each model are summarized in Table 2, and the ROC curves are plotted in Fig. 4. Among the five ML models, random forest showed the highest score in accuracy, and LightGBM showed the highest score in AUC.

Important features of the predictor variables

To detect the importance of each parameter for predicting postoperative rotator cuff re-tear, feature importance was calculated for the LightGBM model, which showed the highest AUC. Age, stump classification, and tear size were ranked as the three most important parameters associated with postoperative rotator cuff re-tear in the LightGBM model (Fig. 5a). The SHAP score showed stump classification, tear size, and age as important characteristics. As shown in Fig. 5b, stump classification and tear size showed a strong positive correlation for postoperative rotator cuff re-tears.

Discussion

The ML classification models predicted re-tears after ARCR with high accuracy. Among the five used models, LightGBM showed the highest AUC. In the LightGBM model, age, stump classification, and tear size were the most important factors affecting rotator cuff re-tear after ARCR.

In the last decades, AI techniques based on mathematical modeling have been developed; ML is one of the AI-based approaches, and ML models are increasingly integrated into clinical diagnosis and the prediction of clinical outcomes. Recently, ML has also been applied to the diagnosis of RCTs, and it has been reported that XGBoost predicts RCTs from clinical findings with high accuracy (accuracy: 0.85, AUC: 0.92).[27] Postoperative re-tear is one of the most important clinical issues associated with RCTs. A variety of risk factors have been reported, including imaging findings such as tear size[5] and fatty degeneration[15] using MRI, as well as patient factors such as age,[2] gender,[8] and preoperative corticosteroid injections.[28] In addition to these risk factors, this study focused on stump classification, which is associated with aging and DM and reflects rotator cuff fragility.[37] The odds ratio (OR) for re-tear risk assessment based on stump classification was 4.71, which was higher than that for tear size (OR: 1.07) and fatty degeneration (OR: 3.87).[39] Therefore, this study added stump classification to the previously described risk factors and presented a comparison of the predictive accuracy of five different learning algorithms. The results showed that all models had high accuracy as classifiers, with LightGBM having the highest AUC. LightGBM is a gradient-boosting framework that uses a decision-tree-based learning algorithm, adopting a histogram algorithm and a depth-limited leaf-wise leaf growth strategy.[41] This strategy increases computational efficiency, reduces memory footprint, improves class classification accuracy, and effectively prevents overfitting.[41] LightGBM has been applied clinically to predict neurological prognosis after cervical cord injury[36] and to predict osteoporosis from blood test data.[22]

In medical AI research, interpretation of model performance is important because clinicians are responsible for making rational decisions based on AI predictions.[27] This concept, called explainable AI (XAI), is intended to enable humans to understand, properly trust, and effectively manage models.[22] In this study, two methods of XAI were used. In the permutation feature method, it is defined as the amount by which the model score decreases when one feature is randomly shuffled. Since the relationship between features and targets is broken in this method, the decrease in model score indicates how dependent the model is on the features.[12] Results indicate that age, stump classification, and tear size are three important parameters. Age is considered to be a strong confounding factor, as it also influences stump classification[37] and tear size.[19] SHAP is another XAI and explains the predictive value of aML model by calculating the contribution of each feature to the prediction. In this model, stump classification, tear size, and age showed higher SHAP scores, all of which were positively correlated with the presence of rotator cuff re-tear. The stump classification reflects the fragility of the tendon,[37] and its recent association with rotator cuff re-tears has attracted much attention, so the AI's decision in this study is reasonable. According to the results of this study, it may be important to include stump classification as a risk factor for rotator cuff re-tear after ARCR. ML-based prediction models are capable of predicting rotator cuff re-tears with high accuracy, and we hope that the addition of stump classification will enable more accurate and convenient prediction of clinical outcomes.

This study has some limitations. First, although the model performed well on the present data set, the number of cases in the original data is not large. Second, we did not consider factors by procedure or surgeon for ARCR to unify the perioperative background. Third, no validation against data from other facilities has been conducted in this study, and a validation study will be needed in the future. Finally, factors predicting rotator cuff re-tear after ARCR surgery in this study did not include evaluation of patient laboratory data or past medical history. The creation of a model based on further data would be the next step to achieving higher prediction accuracy and detecting additional risk factors for rotator cuff re-tear.

Conclusion

The ML classifier model predicted re-tears after ARCR with high accuracy, and the AI model showed that the most important characteristics affecting re-tears were age and imaging findings, including stump classification. Stump classification has been suggested to be related to aging and DM, and a combined evaluation of these factors is necessary to prevent re-tears after ARCR. This model may be able to predict postoperative rotator cuff re-tears based on clinical features.

References

1. Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, et al. Predicting pressure injury in critical care patients: A machine-learning model. Am J Crit Care 2018;27:461-468. doi: 10.4037/ajcc2018525.

2. Barth J, Andrieu K, Fotiadis E, Hannink G, Barthelemy R, Saffarini M. Critical period and risk factors for retear following arthroscopic repair of the rotator cuff. Knee Surg Sports Traumatol Arthrosc 2017;25:2196-2204. doi: 10.1007/s00167-016-4276-x.

3. Cho NS, Lee BG, Rhee YG. Arthroscopic rotator cuff repair using a suture bridge technique: is the repair integrity actually maintained? Am J Sports Med 2011;39:2108-2116. doi: 10.1177/0363546510397171.

4. Cho NS, Rhee YG. The factors affecting the clinical outcome and integrity of arthroscopically repaired rotator cuff tears of the shoulder. Clin Orthop Surg 2009;1:96-104. doi: 10.4055/cios.2009.1.2.96.

5. Cho NS, Yi JW, Lee BG, Rhee YG. Retear patterns after arthroscopic rotator cuff repair: single-row versus suture bridge technique. Am J Sports Med 2010;38:664-671. doi: 10.1177/0363546509350081.

6. Chona DV, Lakomkin N, Lott A, Workman AD, Henry AC, Kuntz AF, et al. The timing of retears after arthroscopic rotator cuff repair. J Shoulder Elbow Surg 2017;26:2054-2059. doi: 10.1016/j.jse.2017.07.015.

7. Chung SW, Kim JY, Kim MH, Kim SH, Oh JH. Arthroscopic repair of massive rotator cuff tears: outcome and analysis of factors associated with healing failure or poor postoperative function. Am J Sports Med 2013;41:1674-1683. doi: 10.1177/0363546513485719.

8. Cofield RH, Parvizi J, Hoffmeyer PJ, Lanzer WL, Ilstrup DM, Rowland CM. Surgical repair of chronic rotator cuff tears. A prospective long-term study. J Bone Joint Surg Am. 2001;83:71-77. doi: 10.2106/00004623-200101000-00010.

9. Collin P, Betz M, Herve A, Walch G, Mansat P, Favard L, et al. Clinical and structural outcome 20 years after repair of massive rotator cuff tears. J Shoulder Elbow Surg 2020;29:521-526. doi: 10.1016/j.jse.2019.07.031.

10. DeOrio JK, Cofield RH. Results of a second attempt at surgical repair of a failed initial rotator-cuff repair. J Bone Joint Surg Am. 1984;66:563-567. doi: 10.2106/00004623-198466040-00011.

11. Diebold G, Lam P, Walton J, Murrell GAC. Relationship between age and rotator cuff retear: A study of 1,600 consecutive rotator cuff repairs. J Bone Joint Surg Am. 2017;99:1198-1205. doi: 10.2106/JBJS.16.00770.

12. Doyen S, Taylor H, Nicholas P, Crawford L, Young I, Sughrue ME. Hollow-tree super: A directional and scalable approach for feature importance in boosted tree models. PLOS ONE 2021;16:e0258658. doi: 10.1371/journal.pone.0258658.

13. Elbuluk AM, Coxe FR, Fabricant PD, Ramos NL, Alaia MJ, Jones KJ. Does medial-row fixation technique affect the retear rate and functional outcomes after double-row transosseous-equivalent rotator cuff repair? Orthop J Sports Med 2019;7:2325967119842881. doi: 10.1177/2325967119842881.

14. Elkins AR, Lam PH, Murrell GAC. Duration of surgery and learning curve affect rotator cuff repair retear rates: A post hoc analysis of 1600 cases. Orthop J Sports Med 2020;8. doi: 10.1177/2325967120954341.

15. Gladstone JN, Bishop JY, Lo IK, Flatow EL. Fatty infiltration and atrophy of the rotator cuff do not improve after rotator cuff repair and correlate with poor functional outcome. Am J Sports Med 2007;35:719-728. doi: 10.1177/0363546506297539.

17. Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. Clin Orthop Relat Res 1994;304:78-83.

16. Goutallier D, Postel JM, Gleyze P, Leguilloux P, Van Driessche S. Influence of cuff muscle fatty degeneration on anatomic and functional outcomes after simple suture of full-thickness tears. J Shoulder Elbow Surg 2003;12:550-554. doi: 10.1016/s1058-2746(03)00211-8.

18. Gowd AK, Agarwalla A, Beck EC, Rosas S, Waterman BR, Romeo AA, et al. Prediction of total healthcare cost following total shoulder arthroplasty utilizing machine learning. J Shoulder Elbow Surg 2022;31:2449-2456. doi: 10.1016/j.jse.2022.07.013.

19. Gumina S, Carbone S, Campagna V, Candela V, Sacchetti FM, Giannicola G. The impact of aging on rotator cuff tear size. Musculoskelet Surg 2013;97;Suppl 1:69-72. doi: 10.1007/s12306-013-0263-2.

20. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. J Big Data 2020;7:94. doi: 10.1186/s40537-020-00369-8.

21. Harris JD. Editorial commentary: personalized hip arthroscopy outcome prediction using machine learning-the future is here. Arthroscopy 2021;37:1498-1502. doi: 10.1016/j.arthro.2021.02.032.

22. Inui A, Nishimoto H, Mifune Y, Yoshikawa T, Shinohara I, Furukawa T, et al. Screening for osteoporosis from blood test data in elderly women using a machine learning approach. Bioengineering (Basel) 2023;10. doi: 10.3390/bioengineering10030277.

23. Ishitani E, Harada N, Sonoda Y, Okada F, Yara T, Katsuki I. Tendon stump type on magnetic resonance imaging is a predictive factor for retear after arthroscopic rotator cuff repair. J Shoulder Elbow Surg 2019;28:1647-1653. doi: 10.1016/j.jse.2019.05.012.

24. Jeong HY, Kim HJ, Jeon YS, Rhee YG. Factors predictive of healing in large rotator cuff tears: is it possible to predict retear preoperatively? Am J Sports Med 2018;46:1693-1700. doi: 10.1177/0363546518762386.

25. Kunze KN, Polce EM, Chahla J. Response to "regarding 'editorial commentary: artificial intelligence in sports medicine diagnosis needs to improve'. Arthroscopy 2021;37:1367-1368. doi: 10.1016/j.arthro.2021.03.012.

26. Kunze KN, Polce EM, Ranawat AS, Randsborg PH, Williams RJ, Allen AA, et al. Application of machine learning algorithms to predict clinically meaningful improvement after arthroscopic anterior cruciate ligament reconstruction. Orthop J Sports Med 2021;9:23259671211046575. doi: 10.1177/23259671211046575.

27. Li C, Alike Y, Hou J, Long Y, Zheng Z, Meng K, et al. Machine learning model successfully identifies important clinical features for predicting outpatients with rotator cuff tears. Knee Surg Sports Traumatol Arthrosc 2023. doi: 10.1007/s00167-022-07298-4.

28. Lubowitz JH, Brand JC, Rossi MJ. Preoperative shoulder corticosteroid injection is associated with revision after primary rotator cuff repair. Arthroscopy 2019;35:693-694. doi: 10.1016/j.arthro.2018.12.025.

29. Malavolta EA, Assunção JH, Ramos FF, Ferreira TC, Gracitelli ME, Bordalo-Rodrigues M, et al. Serial structural MRI evaluation of arthroscopy rotator cuff repair: does Sugaya's classification correlate with the postoperative clinical outcomes? Arch Orthop Trauma Surg 2016;136:791-797. doi: 10.1007/s00402-016-2429-5.

30. Matsuo K, Aihara H, Nakai T, Morishita A, Tohma Y, Kohmura E. Machine learning to predict in-hospital morbidity and mortality after traumatic brain injury. J Neurotrauma 2020;37:202-210. doi: 10.1089/neu.2018.6276.

31. Pedregosa F, Varoquaux G, Al. G. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830.

32. Polce EM, Kunze KN, Fu MC, Garrigues GE, Forsythe B, Nicholson GP, et al. Development of supervised machine learning algorithms for prediction of satisfaction at 2 years following total shoulder arthroplasty. J Shoulder Elbow Surg 2021;30:e290-e299. doi: 10.1016/j.jse.2020.09.007.

33. Quddusi A, Eversdijk HAJ, Klukowska AM, de Wispelaere MP, Kernbach JM, Schröder ML, et al. External validation of a prediction model for pain and functional outcome after elective lumbar spinal fusion. Eur Spine J 2020;29:374-383. doi: 10.1007/s00586-019-06189-6.

34. Ranebo MC, Björnsson Hallgren HC, Holmgren T, Adolfsson LE. Surgery and physiotherapy were both successful in the treatment of small, acute, traumatic rotator cuff tears: a prospective randomized trial. J Shoulder Elbow Surg 2020;29:459-470. doi: 10.1016/j.jse.2019.10.013.

35. Sevinç E. An empowered AdaBoost algorithm implementation: a COVID-19 dataset study. Comput Ind Eng 2022;165:107912. doi: 10.1016/j.cie.2021.107912.

36. Shimizu T, Suda K, Maki S, Koda M, Matsumoto Harmon S, Komatsu M, et al. Efficacy of a machine learning-based approach in predicting neurological prognosis of cervical spinal cord injury patients following urgent surgery within 24 h after injury. J Clin Neurosci 2023;107:150-156. doi: 10.1016/j.jocn.2022.11.003.

37. Shinohara I, Mifune Y, Inui A, Nishimoto H, Yamaura K, Mukohara S, et al. Biochemical markers of aging (advanced glycation end products) and degeneration are increased in Type 3 rotator cuff tendon stumps with increased signal intensity changes on MRI. Am J Sports Med 2022;50:1960-1970. doi: 10.1177/03635465221090649.

38. Sugaya H, Maeda K, Matsuki K, Moriishi J. Functional and structural outcome after arthroscopic full-thickness rotator cuff repair: single-row versus dual-row fixation. Arthroscopy 2005;21:1307-1316. doi: 10.1016/j.arthro.2005.08.011.

39. Takeuchi N, Kozono N, Nishii A, Matsuura K, Ishitani E, Onizuka T, et al. Stump classification was correlated with retear in the suture-bridge and double-row repair techniques for arthroscopic rotator cuff repair. Knee Surg Sports Traumatol Arthrosc 2021;29:2587-2594. doi: 10.1007/s00167-020-06415-5.

40. Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. Crit Care 2020;24:478. doi: 10.1186/s13054-020-03179-9.

41. Yan J, Xu Y, Cheng Q, Jiang S, Wang Q, Xiao Y, et al. LightGBM: accelerated genomically designed crop breeding through ensemble learning. Genome Biol 2021;22:271. doi: 10.1186/s13059-021-02492-y.

42. Zhang C, Lei X, Liu L. Predicting metabolite-disease associations based on LightGBM model. Front Genet 2021;12:660275. doi: 10.3389/fgene.2021.660275.

Table Legends

Table.1 Statistical analysis of patient background in the presence or absence of rotator cuff re-tears. Mean ± standard deviation of each parameter. N.S.: not significant

Table 2. Accuracy and the area under the curve of each ML model in predicting rotator cuff re-tear.

Figure Legends

Fig. 1. Representative magnetic resonance imaging of stump classification. (a) Comparison of signal intensity between deltoid (D; red-circled area) and rotator cuff tears (C; orange-circled area). (b) C<D is classified as type 1, C=D as type 2, and C>D as type 3.

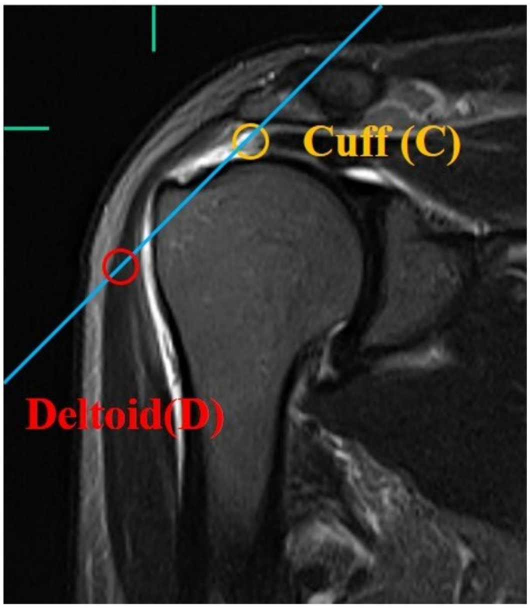
Fig. 2. Workflow of data collection and machine learning.

Fig. 3. Heat map of the correlation. Stump type, diabetes mellitus (DM), tear size, and fatty degeneration positively correlated with rotator cuff re-tear.

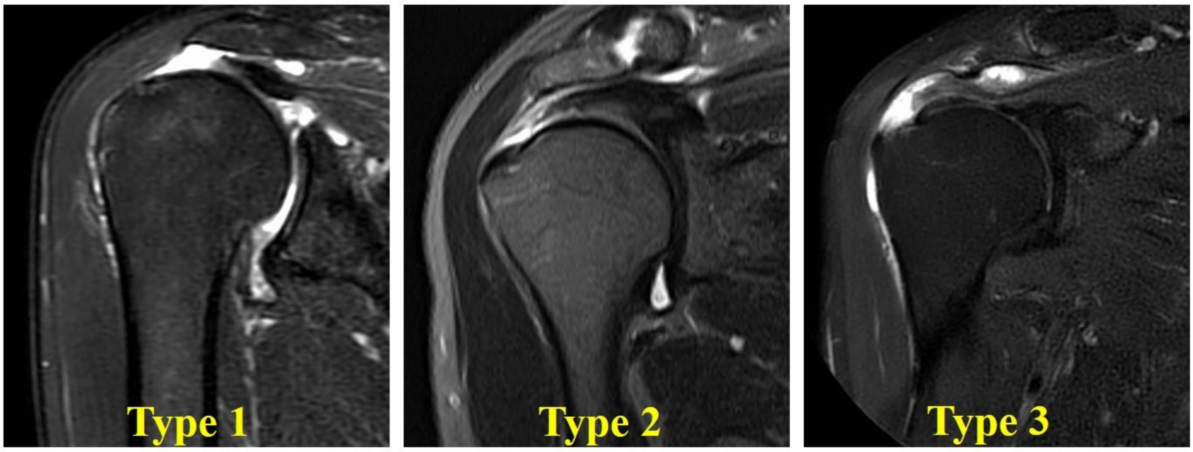
Fig. 4. ROC curve of each trained model

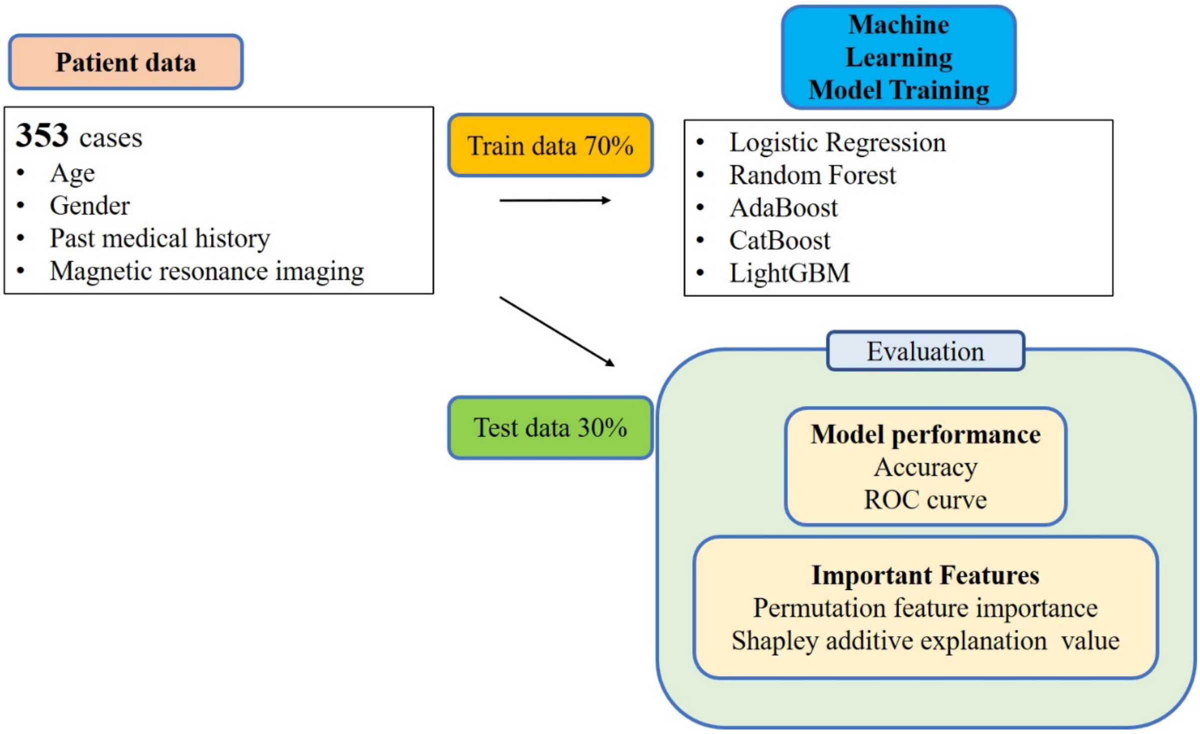
Fig. 5. (a) Permutation features the importance of light Gradient Boosting Machine (LightGBM) model. Important features have larger scores. Top three important features were age, stump type, and tear size. (b) SHAP values of LightGBM model. Top three important features were stump type, tear size, and age. The warm color shows positive impact on model performance while the cool color shows negative impact.

(a)

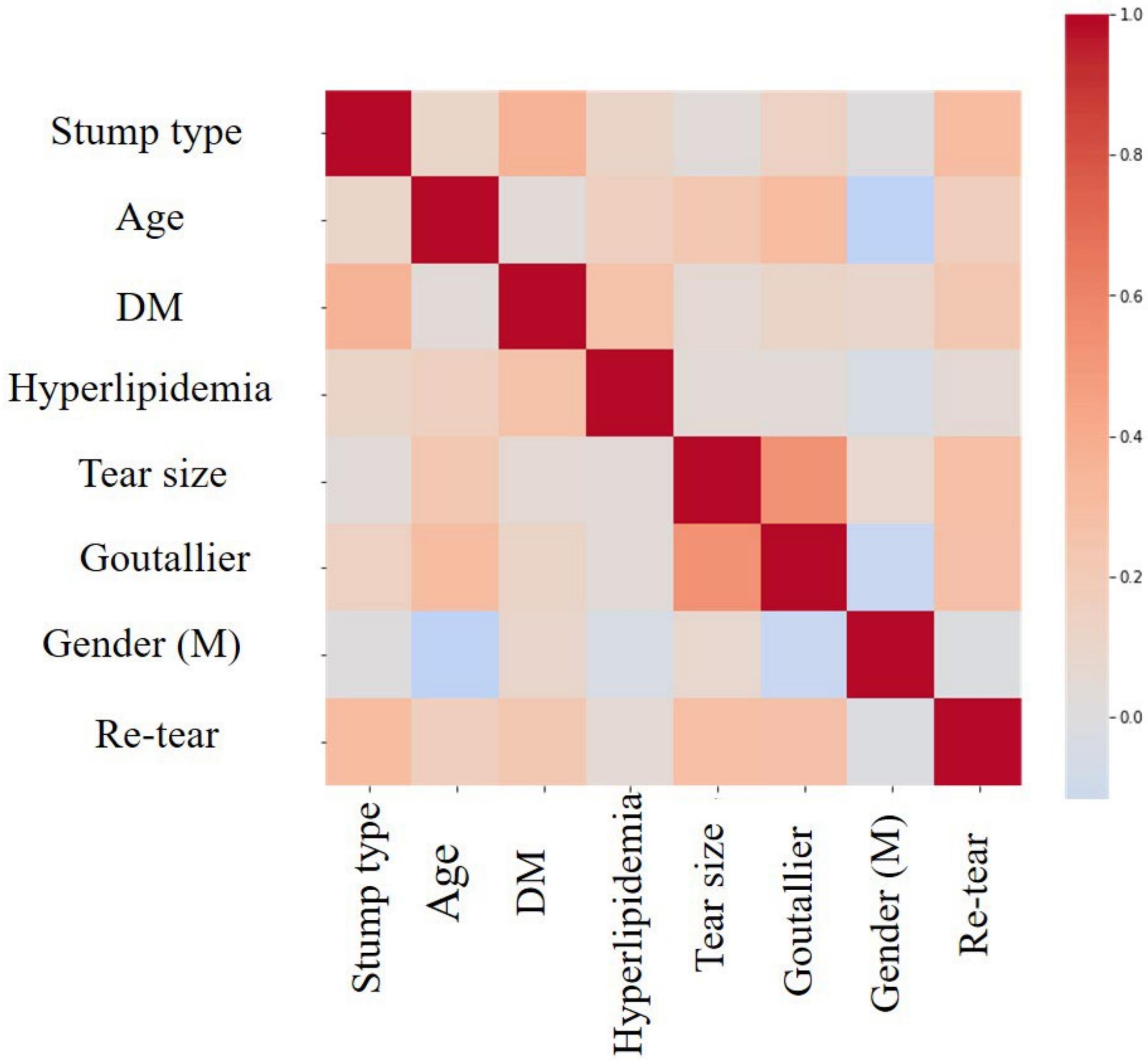


(b)

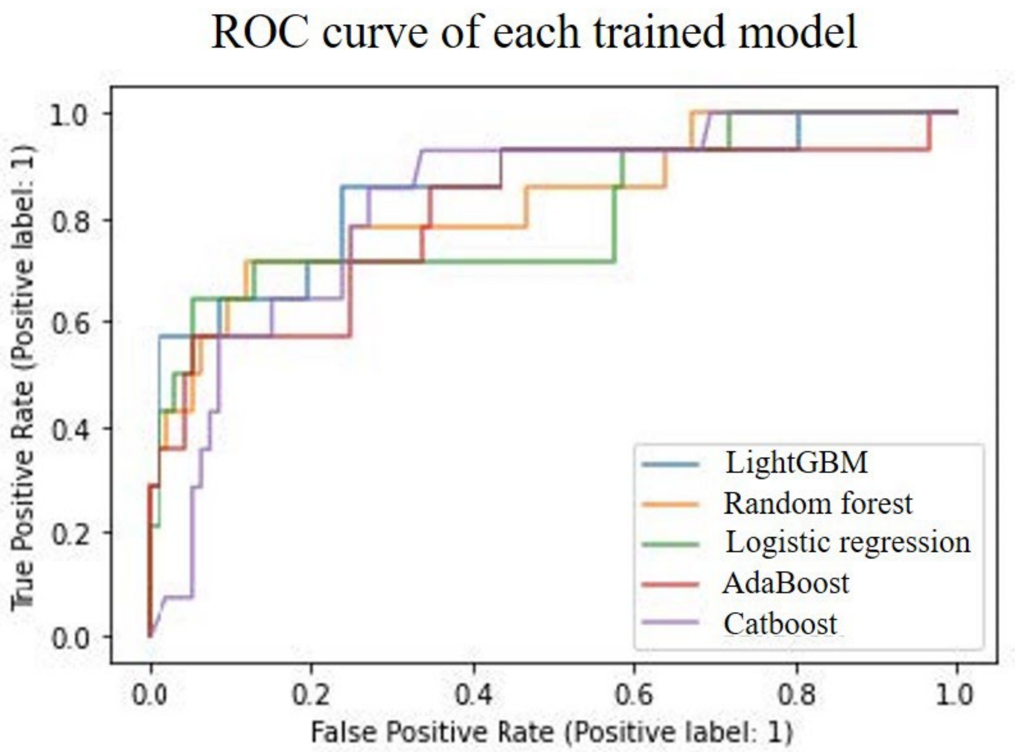




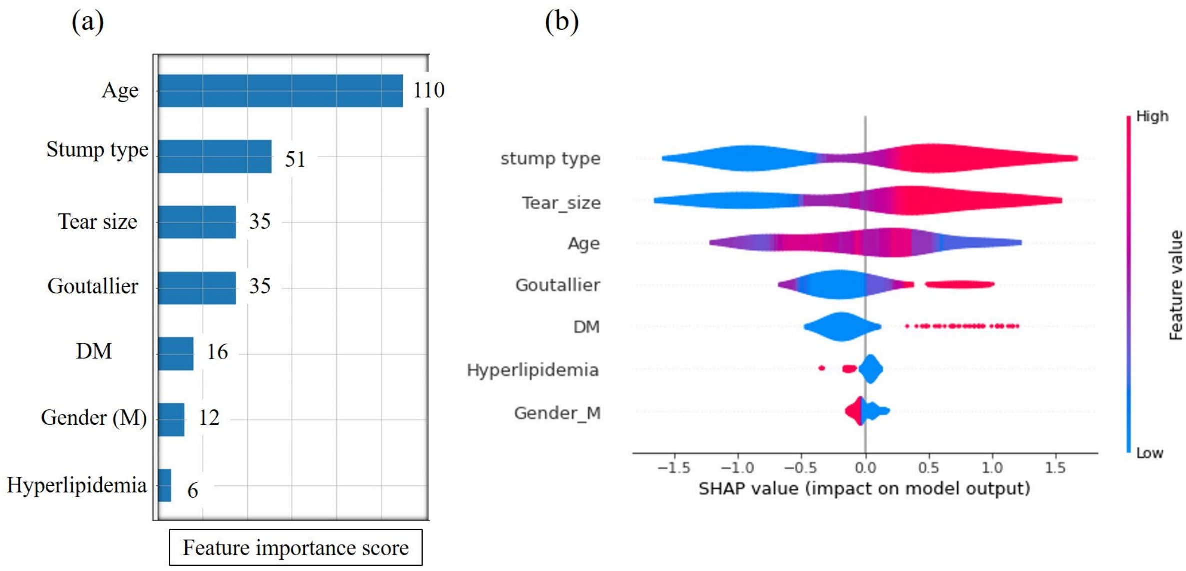
ROC: receiver operating characteristic



DM: diabetes mellitus



ROC: receiver operating characteristic



DM; diabetes mellitus