

Using Machine Learning to Predict Nonachievement of Clinically Significant Outcomes After Rotator Cuff Repair

Rafael Krasic Alaiti,^{*†‡} PhD, Caio Sain Vallio,[§] PhD, Jorge Henrique Assunção,^{||¶} MD, PhD, Fernando Brandão de Andrade e Silva,^{||} MD, PhD, Mauro Emilio Conforto Gracitelli,^{||} MD, PhD, Arnaldo Amado Ferreira Neto,^{||} MD, PhD, and Eduardo Angeli Malavolta,^{||#} MD, PhD

Investigation performed at Faculdade de Medicina, Hospital das Clinicas FMUSP, Universidade de São Paulo, São Paulo, Brazil

Background: Although some evidence suggests that machine learning algorithms may outperform classical statistical methods in prognosis prediction for several orthopaedic surgeries, to our knowledge, no study has yet used machine learning to predict patient-reported outcome measures after rotator cuff repair.

Purpose: To determine whether machine learning algorithms using preoperative data can predict the nonachievement of the minimal clinically important difference (MCID) of disability at 2 years after rotator cuff surgical repair with a similar performance to that of other machine learning studies in the orthopaedic surgery literature.

Study Design: Case-control study; Level of evidence, 3.

Methods: We evaluated 474 patients (n = 500 shoulders) with rotator cuff tears who underwent arthroscopic rotator cuff repair between January 2013 and April 2019. The study outcome was the difference between the preoperative and 24-month postoperative American Shoulder and Elbow Surgeons (ASES) score. A cutoff score was calculated based on the established MCID of 15.2 points to separate success (higher than the cutoff) from failure (lower than the cutoff). Routinely collected imaging, clinical, and demographic data were used to train 8 machine learning algorithms (random forest classifier; light gradient boosting machine [LightGBM]; decision tree classifier; extra trees classifier; logistic regression; extreme gradient boosting [XGBoost]; *k*-nearest neighbors [KNN] classifier; and CatBoost classifier). We used a random sample of 70% of patients to train the algorithms, and 30% were left for performance assessment, simulating new data. The performance of the models was evaluated with the area under the receiver operating characteristic curve (AUC).

Results: The AUCs for all algorithms ranged from 0.58 to 0.68. The random forest classifier and LightGBM presented the highest AUC values (0.68 [95% CI, 0.48-0.79] and 0.67 [95% CI, 0.43-0.75], respectively) of the 8 machine learning algorithms. Most of the machine learning algorithms outperformed logistic regression (AUC, 0.59 [95% CI, 0.48-0.81]); nonetheless, their performance was lower than that of other machine learning studies in the orthopaedic surgery literature.

Conclusion: Machine learning algorithms demonstrated some ability to predict the nonachievement of the MCID on the ASES 2 years after rotator cuff repair surgery.

Keywords: artificial intelligence; machine learning; rotator cuff repair; rotator cuff tears; shoulder pain

Shoulder pain is the third most common musculoskeletal complaint that drives patients to look for health care services, affecting 18% to 26% of adults.³⁴ These conditions are usually labeled as subacromial pain syndrome. They may often be associated with rotator cuff tears or injuries,

which is why those with shoulder pain look for a shoulder specialist.³⁷

There has been an increasing trend in the surgical indication of rotator cuff surgical repairs in those with rotator cuff tears during the last decades,^{7,31,37} generating a high economic burden for treating this population.^{33,37} However, the functional results and retear rate after surgical repair are still disappointing.³⁹ Therefore, although rotator cuff surgical repair has considerable clinical benefits

The Orthopaedic Journal of Sports Medicine, 11(10), 23259671231206180
DOI: 10.1177/23259671231206180
© The Author(s) 2023

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

in symptomatic patients with rotator cuff tears,²² the question of who would likely benefit from surgical treatment remains.

Choosing which surgery may be a good option depends on the interaction of several preoperative information, including^{14,46} the time of conservative treatment without clinical improvement (ie, 6 months) and the belief that rotator cuff tears may worsen and impair patient prognosis in the future, even without evidence of it.⁵ This information is often used by clinicians to decide when to prescribe surgical treatment or continue using conservative approaches. However, these heuristics do not truly assist clinicians during their decision-making process to deliver surgical treatment for those most likely to benefit from it.⁵

Few studies in the literature developed prognostic models to address this issue by identifying prognostic factors for treatment outcomes using classical statistical methods (eg, regression models) for clinical or structural outcomes with large sample sizes.^{12,20,27,36} However, although some models have shown promising results in predicting rotator cuff healing,²⁷ the clinical utility of algorithms developed to predict disability has often been limited, and their performance has been disappointing.

Recently, methods that combine computational science and statistics to maximize the accuracy and predictive power of data, known as machine learning,⁵¹ have begun to be used to predict clinical outcomes in several health conditions.^{9,30,40,45} These studies have shown that machine learning algorithms can predict, with high performance, the achievement of clinically significant outcomes after several orthopaedic surgeries.²⁶ However, to our knowledge, no study has used machine learning to predict patient-reported outcome measures after rotator cuff repair.

The main goal of this study was to determine whether machine learning algorithms using preoperative data can predict the nonachievement of clinically significant disability improvement 2 years after rotator cuff surgical repair with a similar performance to that of other machine learning studies in the literature on orthopaedic surgeries. We hypothesized that machine learning algorithms would predict the nonachievement of the minimal clinically important difference (MCID) with a similar performance to that of other machine learning studies in the literature on orthopaedic surgeries—with the area under the receiver operating characteristic curve (AUC) values of >0.7 .

METHODS

Data Source

We observed a cohort of 474 patients ($n = 500$ shoulders) with subacromial pain syndrome associated with rotator cuff tears who underwent arthroscopic rotator cuff repair between January 2013 and April 2019. The surgical procedures were performed by 4 surgeons in the same institution. The inclusion criteria were as follows: primary arthroscopic rotator cuff repair (partial or complete); having undergone standardized predata collection; and having preoperative magnetic resonance imaging (MRI). Those who had debridement without rotator cuff repair, open or mini-incision surgeries, or previous surgery in the same shoulder were not included. The study protocol received institutional review board approval, and the requirement for informed consent was waived. This study followed the guidelines of the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis.⁴²

Surgery and Rehabilitation

The surgeries were performed in beach-chair or lateral decubitus positions under general anesthesia associated with brachial plexus block, depending on the surgeon's preference. Bursectomy was routinely performed. Depending on the surgeon's preference, acromioplasty was either performed or not. Patients with symptomatic arthrosis in the acromioclavicular joint—defined as pain on local palpation with MRI findings—underwent distal resection. The long head of the biceps was approached when it presented subluxation or dislocation partial lesions $>25\%$ or in the presence of type 2, 3, or 4 superior labrum anterior to posterior lesions.⁴⁹ Tenotomy was performed for patients aged ≥ 60 years. Tenodesis was performed on younger patients, athletes, or those with a body mass index of $>25 \text{ kg/m}^2$, regardless of age.

Rotator cuff repair was performed in single or double rows for posterosuperior tears and single rows for subscapularis tears. Immobilization with a sling was maintained for 4 to 6 weeks. Movements with the elbow, wrist, and fingers were allowed from the first day after surgery. Passive exercises were started after the end of the third week. Active assisted and free active exercises were started after

*Address correspondence to Rafael Krasic Alaiti, PhD, Universidade de São Paulo, Rua Dr. Ovidio Pires de Campos 333, Third Floor, Cerqueira Cesar, São Paulo, SP, Brazil 05403-010 (email: rafael.alaiti@gmail.com).

[†]Research, Technology, and Data Science Office, Grupo Superador, São Paulo, Brazil.

[‡]Universidade de São Paulo, São Paulo, Brazil.

[§]Health Innovation, Data Science, and MLOps, Semantix, São Paulo, Brazil.

^{||}Faculdade de Medicina, Hospital das Clínicas FMUSP, Universidade de São Paulo, São Paulo, Brazil.

[¶]DASA, Hospital 9 de Julho, São Paulo, São Paulo, Brazil.

[#]Hospital do Coração, São Paulo, Brazil.

Final revision submitted May 17, 2023; accepted May 30, 2023.

The authors have declared that there are no conflicts of interest in the authorship and publication of this contribution. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from the Clinical Hospital of the Medical School of the University of São Paulo, São Paulo, Brazil (protocol No. 2.778.930).

the sling was removed. Muscle strengthening was performed only after a significant gain in movement, around the 12th week. Patients were released for sports activities at 6 months, as long as the range of motion and strength were reestablished.

Study Outcome

The outcome of this study—also known as the target variable—was the difference between the preoperative and 24 monthly postoperative scores of the American Shoulder and Elbow Surgeons (ASES) standardized shoulder assessment form.⁴⁷ A cutoff score was calculated based on the established MCID of 15.2 points to separate success (higher than the cutoff) from failure (lower than the cutoff).³⁸

Predictors

Predictors were selected according to already identified risk factors routinely collected at the institution. Demographic characteristics and comorbidity data were collected by a research assistant in the preoperative period using a questionnaire that collected information on patient characteristics, previous illnesses, and life habits, such as smoking.

Predictors with information on the type, location, and extent of shoulder injuries were obtained retrospectively through the analysis of the MRI files together with the radiological report by an attending orthopaedic surgeon (E.A.M.) with >10 years of experience who was blinded as to the surgical procedures and patients' clinical outcomes. In all patients, the interval between the examination and surgery was <12 months.

The following variables were selected from the analyses: (1) supraspinatus tear evaluated according to tendon thickness (partial or full-thickness tear); (2) retraction (<3 or ≥3 cm); (3) extension (affects part of the tendon or the entire extension affected); (4) tear of the anterior portion; (5) infraspinatus tear, categorized into tendon thickness (intact, partial tear, or full-thickness tear), retraction (<3 or ≥3 cm), and extension (intact, superior portion, or the entire tendon); (6) subscapularis tear, categorized as intact, partial tear of the upper third, full-thickness tear of the upper third, or tear involving the upper two-thirds or the entire tendon; (7) fatty degeneration of the supraspinatus; (8) infraspinatus and subscapularis muscles (according to Goutallier et al¹⁷ and modified by Fuchs et al¹³); (9) long head of the biceps tear (stable, partial tear, complete tear) and instability (topical, subluxated, dislocated, depending on its position in the biceps sulcus, or not applicable in cases of complete tear); and (10) the presence of arthrosis of the glenohumeral joint (absent or present). Symptomatic acromioclavicular arthrosis was defined as pain on local palpation with MRI findings, such as capsuloligamentous thickening and osteophytosis.

Statistical Analysis

We used a random sample of 70% of patients to train the algorithms, and 30% were left for performance assessment,

simulating new data. Stratified cross-validation with 10 folds was used to train the models and adjust hyperparameters. The predictors were normalized by the z score. The strategy of the last observation carried forward was adopted, with 12-month outcome data when there were missing outcome variable values at 24 months. Patients with missing data at 24 and 12 months were excluded from the analysis.

The borderline Synthetic Minority Oversampling Technique (SMOTE) was used to balance classes of the target variable.¹⁹ Automated methods of selection of variables were not used because they were chosen on theoretical basis from experts; however, a multicollinearity test was performed, and any variable with >0.9 of correlation was removed.⁶ Hyperparameter optimization of all algorithms, including logistic regression, was performed to optimize the AUC by the Optuna library,¹ with the tree-structured Parzen estimator⁴ as the search algorithm and asynchronous successive halving algorithm³² as the early stopping algorithm. We applied the same preprocessing steps and feature selection techniques to all algorithms in the study to ensure a fair comparison. The AUC was used to evaluate the performance of the models. We also extracted the accuracy, precision, recall, and F1 score to evaluate the models. The 95% bootstrap confidence interval was calculated to assess the variability of these metrics. To interpret the final model, the SHAP (SHapley Additive exPlanations) was used to understand the influence of the variables.³⁵

After removing highly correlated variables ($r > 0.9$) and identifying variables (eg, patient number), 24 variables were selected to develop the predictive models. The algorithms used were as follows: random forest classifier, light gradient boosting machine (LightGBM), decision tree classifier, extra trees classifier, extreme gradient boosting (XGBoost), k -nearest neighbors (KNN) classifier, CatBoost classifier, and logistic regression.

RESULTS

During the period evaluated, 651 surgeries were performed for rotator cuff repair. The following were excluded: 84 open procedures, 10 debridement surgeries, 26 cases with previous shoulder surgery, 12 patients without postoperative clinical evaluation, and 19 patients with incomplete pre- or perioperative evaluation data. The analyzed sample consisted of 474 patients ($n = 500$ shoulders). Data imputation from outcomes at 12 months was necessary in the functional assessment in 76 cases (15.2%) because of missing data at 24 months. Therefore, 474 patients were included in the study and further divided into train and test datasets. The preoperative predictors and their summarized values can be seen in Tables 1 to 3. As the target variable was highly imbalanced—that is, 17.2% of participants did not reach the minimum MCID in 24 months after surgery—it was further corrected with the borderline SMOTE technique.¹⁹

The most common models for prediction with structured data were then fitted. Table 4 presents the performance

TABLE 1
Demographic and Comorbidity Predictors of Patients
Included in the Models (N = 474)^a

Predictor	Value
Age, y	57.8 ± 8.9
Sex	
Male	230 (46)
Female	270 (54)
Diabetes	75 (15)
Rheumatoid arthritis	28 (5.6)
Smoker	
Current	53 (10.6)
Past	79 (15.8)
Traumatic lesion	71 (14.2)
Worker's compensation	64 (12.8)
Previous injections	71 (14.2)
Preoperative ASES score	41.6 ± 19.2

^aData are presented as mean ± SD or frequency (%). ASES, American Shoulder and Elbow Surgeons.

TABLE 2
Supraspinatus and Infraspinatus MRI Predictors
Included in the Model^a

Predictor	Value
Supraspinatus	
Partial tear	72 (14.4)
Full-thickness tear	428 (85.6)
Retraction <3 cm	339 (67.8)
Retraction ≥3 cm	161 (32.2)
Entire extension affected	224 (44.8)
Anterior portion affected	409 (81.8)
Fatty degeneration: Goutallier grades 0 and 1	394 (78.8)
Fatty degeneration: Goutallier grade 2	77 (15.4)
Fatty degeneration: Goutallier grades 3 and 4	29 (5.8)
Infraspinatus	
Partial tear	36 (7.2)
Full-thickness tear	101 (20.2)
Retraction <3 cm	452 (90.4)
Retraction ≥3 cm	48 (9.6)
Superior portion affected	115 (23)
Entire tendon affected	22 (4.4)
Fatty degeneration: Goutallier grades 0 and 1	423 (84.6)
Fatty degeneration: Goutallier grade 2	53 (10.6)
Fatty Degeneration: Goutallier grades 3 and 4	24 (4.8)

^aData are presented as frequency (%). MRI, magnetic resonance imaging.

and variability (ie, 95% bootstrap CIs) of the models in the test dataset, ordered by the AUC metric. The model with the best AUC metric was the Random Forest Classifier (0.68 [95% CI, 0.48-0.79]). An interface with the trained random forest classifier algorithm deployed was developed for clinical application using Streamlit (<https://bit.ly/rotatorcuffsurgeryAI>).

The SHAP was used to interpret the relationship of preoperative predictors with the model outcome that presented the best performance (Figure 1). The most

TABLE 3
Subscapularis, Long Head of the Biceps, and Glenohumeral
Joint MRI Predictors Included in the Model^a

Predictor	Value
Subscapularis	
Intact tendon	260 (52)
Partial tear of the upper third	159 (31.8)
Full-thickness tear of the upper third	51 (10.2)
Full-thickness tear of the upper two-thirds or more	30 (6)
Fatty degeneration: Goutallier grades 0 and 1	403 (80.6)
Fatty degeneration: Goutallier grade 2	34 (6.8)
Fatty degeneration: Goutallier grades 3 and 4	7 (1.4)
Long head of the biceps	
Intact tendon	403 (80.6)
Partial tear	66 (13.2)
Complete tear	31 (6.2)
Stable	358 (71.6)
Subluxated	71 (12.2)
Dislocated	40 (8)
Not applicable (complete tear)	31 (6.2)
Glenohumeral joint	
Glenohumeral arthritis	39 (7.8)

^aData are presented as frequency (%). MRI, magnetic resonance imaging.

relevant predictor was the extent of supraspinatus affected, in which patients with complete supraspinatus involvement were more likely to not reach the MCID 24 months after surgery. The second most relevant predictor was the preoperative ASES score, in which the higher the preoperative score, the greater the chance of not reaching the MCID 24 months after surgery. The third most relevant predictor was the presence of Goutallier grade ≥2 fatty infiltration of the supraspinatus tendon, in which patients without it were more likely not to reach the MCID. The fourth most relevant predictor was age, in which younger patients were more likely to not reach the MCID. The interpretation of the other models can be found in the Supplemental Material, available separately.

DISCUSSION

We found that machine learning algorithms could predict the nonachievement of the disability MCID 2 years after a rotator cuff repair surgery. Still, their performance was lower than that of other machine learning studies in the orthopaedic surgery literature.²⁶ The mean AUC of all algorithms was 0.62 in the test set, of which the Random Forest Classifier and LightGBM presented the highest measures (0.68 and 0.67, respectively). In addition, the mean recall and precision were relatively low (0.41 and 0.25, respectively).

A recent systematic review has identified 18 studies that used machine learning algorithms to predict MCIDs after several orthopaedic surgeries, of which 7 concerned outcomes after spine surgery, 6 after sports medicine procedures, 3 after total joint arthroplasty (hip and knee),

TABLE 4
Performance Measures of the Prediction Models^a

Model	Accuracy	AUC	Recall	Precision	F1 Score
Random forest classifier	0.78 (0.69-0.88)	0.68 (0.48-0.79)	0.41 (0.12-0.73)	0.35 (0.10-0.65)	0.38 (0.11-0.60)
LightGBM	0.68 (0.58-0.78)	0.67 (0.43-0.75)	0.45 (0.16-0.75)	0.24 (0.08-0.42)	0.34 (0.11-0.51)
Decision tree classifier	0.75 (0.65-0.85)	0.62 (0.41-0.64)	0.17 (0.01-0.40)	0.19 (0.01-0.45)	0.18 (0.01-0.38)
Extra trees classifier	0.73 (0.63-0.83)	0.60 (0.45-0.75)	0.37 (0.11-0.69)	0.26 (0.07-0.50)	0.31 (0.08-0.51)
Logistic regression	0.66 (0.54-0.77)	0.59 (0.48-0.81)	0.62 (0.33-0.90)	0.26 (0.11-0.42)	0.37 (0.17-0.54)
XGBoost	0.60 (0.49-0.72)	0.60 (0.39-0.72)	0.50 (0.20-0.80)	0.20 (0.06-0.35)	0.29 (0.10-0.45)
KNN classifier	0.79 (0.69-0.88)	0.58 (0.44-0.70)	0.25 (0.01-0.50)	0.30 (0.01-0.60)	0.27 (0.01-0.50)
CatBoost classifier	0.58 (0.48-0.69)	0.59 (0.41-0.73)	0.54 (0.25-0.83)	0.20 (0.07-0.35)	0.29 (0.11-0.47)

^aData in parentheses are 95% bootstrap CIs. AUC, area under the receiver operating characteristic curve; KNN, *k*-nearest neighbors; LightGBM, light gradient-boosting machine; XGBoost, extreme gradient boosting.

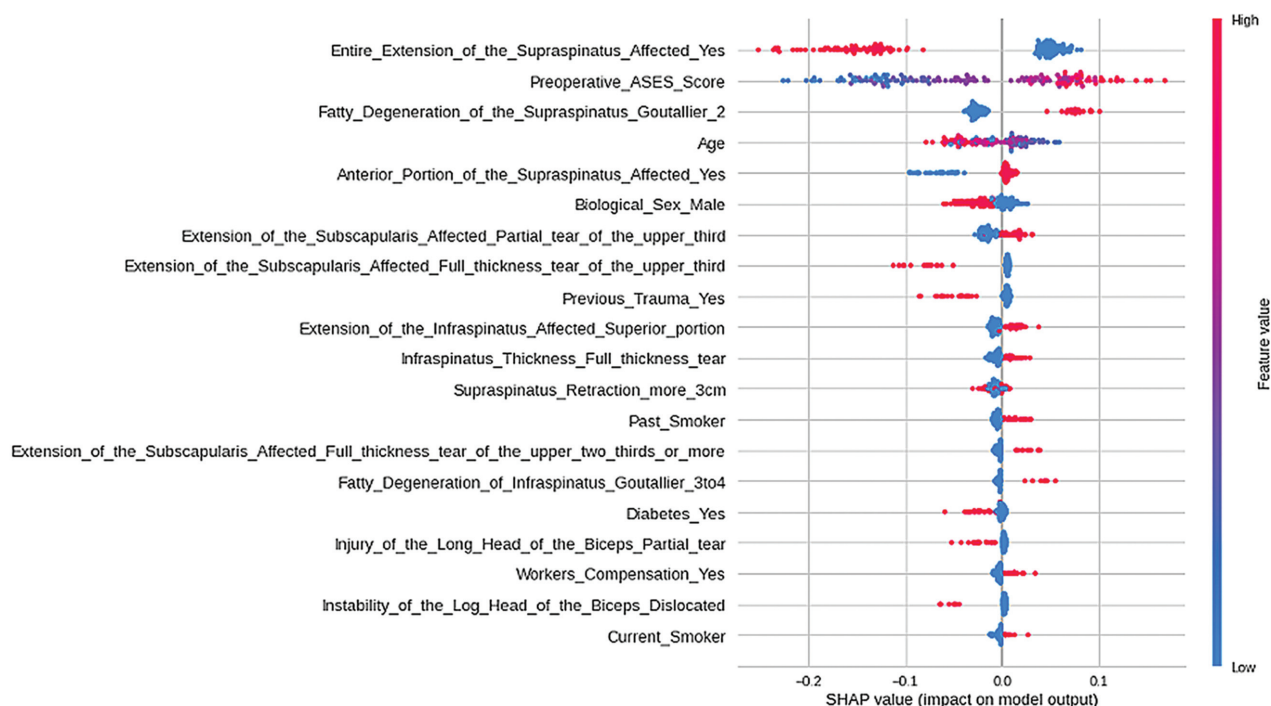


Figure 1. SHAP values of the random forest classifier model. This figure provides other relevant information for model interpretation: (1) the predictors are ordered from top to bottom according to their relevance; (2) the more to the right the points of a variable are, the greater the influence of the variable in predicting the outcome (ie, not reaching the minimum MCID); and (3) the redder the point, the higher the predictor value; and the bluer the point, the lower the predictor value. ASES, American Shoulder and Elbow Surgeons; MCID, minimal clinically important difference.

and 2 after shoulder arthroplasty.²⁶ Most studies performed well in predicting the MCID, with AUCs ranging from 0.7 to 0.9.

Although we were not able to find other studies that developed machine learning algorithms to predict outcomes after rotator cuff repair surgery, 2 studies used machine learning to predict MCID achievement after primary shoulder arthroplasty. In both studies, the patients were separated into cohorts undergoing anatomic shoulder arthroplasty and reverse arthroplasty. Kumar et al^{24,25}

used the XGBoost classifier to accurately predict MCID achievement, with AUCs between 0.7 and 0.98.

While the random forest classifier model was 8 points ahead in the AUC compared with the logistic regression—a model commonly used in health care for predictions—the more complex ensemble models (ie, LightGBM, XGBoost, CatBoost) did not outperform random forest. One possible explanation for this is that the ensemble models perform better with massive quantities of data, and the present study used a few variables and

people to train the algorithm. Additionally, the KNN and CatBoost classifier algorithms did not outperform logistic regression.

Therefore, although the development of accurate prognostic models for orthopaedic surgeries using machine learning algorithms is undoubtedly a natural next step, given that artificial intelligence has shown promising results in predicting outcomes in various scientific fields, including health care, better than classical statistical methods,^{21,44,48} the overall performance of the algorithms developed in this study was moderate.

The mean low recall and precision values as well as the inability to achieve higher AUC values in our models may be related to a few limitations of this study that should be noted. The analyses conducted during this study were not planned. Therefore, some relevant preoperative predictors of outcomes after rotator cuff repair surgery may not have been included in the analysis. While several studies suggest that the presence of systemic diseases (eg, diabetes) and the severity of rotator cuff tear or injury are relevant prognostic factors for rotator cuff repair outcomes,^{11,16,18,28,31,43} evidence suggests that socioeconomic and psychosocial factors might also provide relevant information about rotator cuff repair outcomes.⁵⁰

The use of socioeconomic and psychosocial variables as predictors of outcomes after surgeries is supported by studies identifying that these variables are relevant prognostic factors for failed back surgery syndrome and chronic post-surgical pain in several conditions.^{3,8,15,23,29,41} Finally, although the sample size used in this study is, to our knowledge, one of the highest analyzing clinical outcomes of rotator cuff repair literature, it is important to note that training our machine learning models with a dataset of only 350 patients might limit their utility. Machine learning models typically benefit from larger training datasets, and their performance would likely improve significantly with an increase in training data in the future. Therefore, future studies should also use higher sample sizes and preoperative measures of socioeconomic status and psychosocial factors as potential predictors for this population to develop better prediction models. Moreover, the strategy for outcome missing data (ie, using 12 months assessment when 24 months were not available) might have impacted the analysis, despite the mean differences in the ASES score in 12 and 24 months being relatively close.

It is important to note that the MCID of 15.2 points used for the ASES scale in our study was based on the mean found among 3 methods to determine the MCID in a study by Malavolta et al³⁸ (ie, distribution, anchor, and minimum detectable change). Those authors observed values ranging from 6.1 (anchor method) to 26.3 (minimum detectable change). Therefore, different MCID values can influence the predictors' weight in further models. Additionally, patients with preoperative ASES scores >84.8 (ie, which corresponds to 2% of our dataset) may not be able to achieve the MCID used in this study, which is higher than their potential improvement. This limitation should also be considered when interpreting the results.

Furthermore, the structural outcome of the rotator cuff repair was not evaluated in the present study since it is not part of our routine to perform postoperative MRIs to assess tendon healing. Although the integrity of rotator cuff tendons does not seem to be directly associated with functional outcomes,^{2,10} future studies might also benefit clinicians during their decision-making with models to predict retears after surgeries.

Clinical Implications

Supervised machine learning algorithms are primarily designed to learn from hidden patterns in available data about maximizing outcome prediction, rather than explaining causal relationships between a prediction and the outcome. This is because they adjust the weight of each variable based on the hyperparameters set and treat each category of categorical variables as a variable by itself, which makes it difficult to interpret the individual role of each predictor variable. This is also one of the reasons why machine learning algorithms may outperform conventional statistical methods and linear thinking. Nevertheless, the results of this study suggest that the use of machine learning algorithms might be a promising new tool that can assist clinicians during clinical decision-making to decide when to prescribe surgical treatment or continue using nonoperative approaches to treat patients with rotator cuff tears.

CONCLUSION

We found that machine learning algorithms demonstrated some ability to predict the nonachievement of the disability MCID 2 years after a rotator cuff repair surgery. Still, their performance was lower than that of other machine learning studies in the literature on orthopaedic surgeries.

Supplemental Material for this article is available at <https://journals.sagepub.com/doi/full/10.1177/23259671231206180#supplementary-materials>

REFERENCES

1. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. KDD '19. Association for Computing Machinery; 2019:2623-2631.
2. Alaiti RK, Caneiro JP, Gasparin JT, et al. Shoulder pain across more movements is not related to more rotator cuff tendon findings in people with chronic shoulder pain diagnosed with subacromial pain syndrome. *Pain Rep*. 2021;6(4):e980.
3. Archer KR, Wegener ST, Seebach C, et al. The effect of fear of movement beliefs on pain and disability after surgery for lumbar and cervical degenerative conditions. 2011;36(19):1554-1562.

4. Bergstra J, Bardenet R, Kégl B, Bengio Y. Algorithms for hyperparameter optimization. In: *Advances in Neural Information Processing Systems*. unknown; 2011.
5. Boorman RS, More KD, Hollinshead RM, et al. What happens to patients when we do not repair their cuff tears? Five-year rotator cuff quality-of-life index outcomes following nonoperative treatment of patients with full-thickness rotator cuff tears. *J Shoulder Elbow Surg*. 2018;27(3):444-448.
6. Chan JYL, Leow SMH, Bea KT, et al. Mitigating the multicollinearity problem and its machine learning approach: a review. *Mathematics*. 2022;10(8):1283.
7. Colvin AC, Egorova N, Harrison AK, Moskowitz A, Flatow EL. National trends in rotator cuff repair. *J Bone Joint Surg Am*. 2012;94(3):227-233.
8. Dabija DI, Pennings JS, Archer KR, et al. Which is the best outcome measure for rotator cuff tears? *Clin Orthop Relat Res*. 2019; 477(8):1869-1878.
9. de Toledo P, Rios PM, Ledezma A, Sanchis A, Alen JF, Lagares A. Predicting the outcome of patients with subarachnoid hemorrhage using machine learning techniques. *IEEE Trans Inf Technol Biomed*. 2009;13(5):794-801.
10. Dunn WR, Kuhn JE, Sanders R, et al. Symptoms of pain do not correlate with rotator cuff tear severity: a cross-sectional study of 393 patients with a symptomatic atraumatic full-thickness rotator cuff tear. *J Bone Joint Surg Am*. 2014;96(10):793-800.
11. Flurin PH, Marczuk Y, Janout M, Wright TW, Zuckerman J, Roche CP. Comparison of outcomes using anatomic and reverse total shoulder arthroplasty. *Bull Hosp Jt Dis*. 2013;71 (suppl 2):101-107.
12. Frangiamore S, Dornan GJ, Horan MP, et al. Predictive modeling to determine functional outcomes after arthroscopic rotator cuff repair. *Am J Sports Med*. 2020;48(7):1559-1567.
13. Fuchs B, Weishaupt D, Zanetti M, Hodler J, Gerber C. Fatty degeneration of the muscles of the rotator cuff: assessment by computed tomography versus magnetic resonance imaging. *J Shoulder Elbow Surg*. 1999;8(6):599-605.
14. Ghodadra NS, Provencher MT, Verma NN, Wilk KE, Romeo AA. Open, mini-open, and all-arthroscopic rotator cuff repair surgery: indications and implications for rehabilitation. *J Orthop Sports Phys Ther*. 2009;39(2):81-89.
15. Giusti I, Di Francesco M, D'Ascenzo S, et al. Leukocyte depletion does not affect the in vitro healing ability of platelet rich plasma. *Exp Ther Med*. 2018;15(4):4029-4038.
16. Gladstone JN, Bishop JY, Lo IKY, 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(5):719-728.
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.
18. Gulotta LV, Nho SJ, Dodson CC, et al. Prospective evaluation of arthroscopic rotator cuff repairs at 5 years: part II—prognostic factors for clinical and radiographic outcomes. *J Shoulder Elbow Surg*. 2011;20(6):941-946.
19. Han H, Wang WY, Mao BH. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In: *Advances in Intelligent Computing*. Springer Berlin Heidelberg; 2005:878-887.
20. Jenssen KK, Lundgreen K, Madsen JE, Kvakestad R, Dimmen S. Prognostic factors for functional outcome after rotator cuff repair: a prospective cohort study with 2-year follow-up. *Am J Sports Med*. 2018;46(14):3463-3470.
21. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-243.
22. Karjalainen TV, Jain NB, Heikkinen J, Johnston RV, Page CM, Buchbinder R. Surgery for rotator cuff tears. *Cochrane Database Syst Rev*. 2019;12(12):CD013502.
23. Kim HJ, Park JW, Chang BS, Lee CK, Yeom JS. The influence of cat-astrophising on treatment outcomes after surgery for lumbar spinal stenosis. *Bone Joint J*. 2015;97(11):1546-1554.
24. Kumar V, Roche C, Overman S, et al. Using machine learning to predict clinical outcomes after shoulder arthroplasty with a minimal feature set. *J Shoulder Elbow Surg*. 2021;30(5):e225-e236.
25. Kumar V, Roche C, Overman S, et al. What is the accuracy of three different machine learning techniques to predict clinical outcomes after shoulder arthroplasty? *Clin Orthop Relat Res*. 2020;478(10): 2351-2363.
26. Kunze KN, Krivich LM, Clapp IM, et al. Machine learning algorithms predict achievement of clinically significant outcomes after orthopaedic surgery: a systematic review. *Arthroscopy*. 2022;38(6):2090-2105.
27. Kwon J, Kim SH, Lee YH, Kim TI, Oh JH. The rotator cuff healing index: a new scoring system to predict rotator cuff healing after surgical repair. *Am J Sports Med*. 2019;47(1):173-180.
28. Lapner PLC, Sabri E, Rakhra K, et al. A multicenter randomized controlled trial comparing single-row with double-row fixation in arthroscopic rotator cuff repair. *J Bone Joint Surg Am*. 2012;94(14):1249-1257.
29. Larsen DB, Laursen M, Edwards RR, Simonsen O, Arendt-Nielsen L, Petersen KK. The combination of preoperative pain, conditioned pain modulation, and pain catastrophizing predicts postoperative pain 12 months after total knee arthroplasty. *Pain Med*. 2021;22(7):1583-1590.
30. Lee SI, Celik S, Logsdon BA, et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat Commun*. 2018;9(1):42.
31. Lichtenberg S, Liem D, Magosch P, Habermeyer P. Influence of tendon healing after arthroscopic rotator cuff repair on clinical outcome using single-row Mason-Allen suture technique: a prospective, MRI controlled study. *Knee Surg Sports Traumatol Arthrosc*. 2006;14(11): 1200-1206.
32. Li L, Jamieson K, Rostamizadeh A, et al. A system for massively parallel hyperparameter tuning. *Proceedings of Machine Learning and Systems*. 2020;2:230-246.
33. Longo UG, Salvatore G, Rizzello G, et al. The burden of rotator cuff surgery in Italy: a nationwide registry study. *Arch Orthop Trauma Surg*. 2017;137(2):217-224.
34. Luime JJ, Koes BW, Hendriksen IJM, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scand J Rheumatol*. 2004;33(2):73-81.
35. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017;30.
36. Malavolta EA, Assunção JH, Andrade-Silva FB, Gracitelli MEC, Kiyomoto HD, Aa FN. Prognostic factors for clinical outcomes after arthroscopic rotator cuff repair. *Orthop J Sports Med*. 2023; 12;11(4):23259671231160738.
37. Malavolta EA, Assunção JH, Beraldo RA, Pinto G de MR, Gracitelli MEC, Ferreira Neto AA. Rotator cuff repair in the Brazilian Unified Health System: Brazilian trends from 2003 to 2015. *Rev Bras Ortop*. 2017;52(4):501-505.
38. Malavolta E, Yamamoto G, Bussius G, et al. Establishing minimal clinically important difference for the UCLA and ASES scores after rotator cuff repair. *Orthop Traumatol Surg Res*. 2022;108(2):102894.
39. McElvany MD, McGoldrick E, Gee AO, Neradilek MB, Matsen FA III. Rotator cuff repair: published evidence on factors associated with repair integrity and clinical outcome. *Am J Sports Med*. 2015; 43(2):491-500.
40. Merali ZG, Witiw CD, Badhiwala JH, Wilson JR, Fehlings MG. Using a machine learning approach to predict outcome after surgery for degenerative cervical myelopathy. *PLoS One*. 2019;14(4):e0215133.
41. Montes A, Roca G, Cantillo J, Sabate S, GENDOLCAT Study Group. Presurgical risk model for chronic postsurgical pain based on 6 clinical predictors: a prospective external validation. *Pain*. 2020; 161(11):2611-2618.
42. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015; 162(1):W1-w73.

43. Nho SJ, Brown BS, Lyman S, Adler RS, Altchek DW, MacGillivray JD. Prospective analysis of arthroscopic rotator cuff repair: prognostic factors affecting clinical and ultrasound outcome. *J Shoulder Elbow Surg.* 2009;18(1):13-20.
44. Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V. Artificial intelligence transforms the future of health care. *Am J Med.* 2019;132(7):795-801.
45. Oermann EK, Rubinsteyn A, Ding D, et al. Using a machine learning approach to predict outcomes after radiosurgery for cerebral arteriovenous malformations. *Sci Rep.* 2016;6:21161.
46. Raman J, Walton D, MacDermid JC, Athwal GS. Predictors of outcomes after rotator cuff repair—a meta-analysis. *J Hand Ther.* 2017;30(3):276-292.
47. Richards RR, An KN, Bigliani LU, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg.* 1994;3(6):347-352.
48. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet.* 2020;395(10236):1579-1586.
49. Snyder SJ, Banas MP, Karzel RP. An analysis of 140 injuries to the superior glenoid labrum. *J Shoulder Elbow Surg.* 1995;4(4):243-248.
50. Thorpe AM, O'Sullivan PB, Mitchell T, et al. Are psychologic factors associated with shoulder scores after rotator cuff surgery? *Clin Orthop Relat Res.* 2018;476(10):2062-2073.
51. Wilkinson J, Arnold KF, Murray EJ, et al. Time to reality check the promises of machine learning-powered precision medicine. *Lancet Digit Health.* 2020;2(12):e677-e680.