A Machine Learning Model Demonstrates Excellent Performance in Predicting Subscapularis Tears Based on Pre-Operative Imaging Parameters Alone

Jacob F. Oeding, M.S., Ayoosh Pareek, M.D., Micah J. Nieboer, M.D., Nicholas G. Rhodes, M.D., Christin A. Tiegs-Heiden, M.D., Christopher L. Camp, M.D., R. Kyle Martin, M.D., F.R.C.S.C., Gilbert Moatshe, M.D., Ph.D., Lars Engebretsen, M.D., Ph.D., and Joaquin Sanchez-Sotelo, M.D., Ph.D.

Purpose: To develop a machine learning model capable of identifying subscapularis tears before surgery based on imaging and physical examination findings. **Methods:** Between 2010 and 2020, 202 consecutive shoulders underwent arthroscopic rotator cuff repair by a single surgeon. Patient demographics, physical examination findings (including range of motion, weakness with internal rotation, lift/push-off test, belly press test, and bear hug test), and imaging (including direct and indirect signs of tearing, biceps status, fatty atrophy, cystic changes, and other similar findings) were included for model creation. **Results:** Sixty percent of the shoulders had partial or full thickness tears of the subscapularis verified during surgery (83% of these were upper third). Using only preoperative imaging-related parameters, the XGBoost model demonstrated excellent performance at predicting subscapularis tears (c-statistic, 0.84; accuracy, 0.85; F1 score, 0.87). The top 5 features included direct signs related to the presence of tearing as evidenced on magnetic resonance imaging (MRI) (changes in tendon morphology and signal), as well as the quality of the MRI and biceps pathology. **Conclusions:** In this study, machine learning was successful in predicting subscapularis tears by MRI alone in 85% of patients, and this accuracy did not decrease by isolating the model to the top features. The top five features included direct signs related to the presence of tearing as evidenced on MRI (changes in tendon morphology and signal), as well as the quality of the MRI and biceps pathology. Last, in advanced modeling, the addition of physical examination or patient characteristics did not make a significant difference in the predictive ability of this model. **Level of Evidence:** Level III, diagnostic case-control study.

See commentary on page 1056

From the School of Medicine, Mayo Clinic Alix School of Medicine (J.F.O.); the Department of Orthopedic Surgery and Sports Medicine, Mayo Clinic (M.J.N., C.L.C., J.S-S.); and the Department of Radiology, Mayo Clinic (N.G.R., C.A.T-H.), Rochester; the Department of Orthopedic Surgery, University of Minnesota (R.K.M.), Minneapolis, Minnesota, U.S.A.; the Sports Medicine and Shoulder Service, Hospital for Special Surgery (A.P.), New York, New York, U.S.A.; and the Oslo Sports Trauma Research Center, Norwegian School of Sport Sciences (J.F.O., A.P., R.K.M., G.M., L.E.), Oslo, Norway.

The authors report the following potential conflicts of interest or sources of funding: J.S-S. reports other from Acumed, LLC: Paid consultant; American Shoulder and Elbow Surgeons: Board or committee member; Elsevier: Publishing royalties, financial or material support; Exactech, Inc: Paid consultant; Journal of Shoulder and Elbow Surgery: Editorial or governing board; Publishing royalties, financial or material support; Oxford University Press: Publishing royalties, financial or material support; Precision OS: Stock or stock Options; PSI: Stock or stock Options; Stryker: IP royalties; Paid presenter or speaker; Research support. G.M. reports other from Arthroscopy: Editorial or governing board; International Society of Arthroscopy, Knee Surgery, and Orthopaedic Sports Medicine: Board or committee member; Journal of Bone and Joint Surgery—American: Editorial or governing board; Smith & Nephew: Paid consultant. C.L.C. reports other from Arthrex, Inc: IP royalties; Major League Baseball: Research support; Arthrex, Inc: IP royalties. L.E. reports other from Acta Orthopaedica: Editorial or governing board; American Orthopaedic Society for Sports Medicine: Board or committee member; Arthrex, Inc: IP royalties; Paid consultant; Research support; Biomet: Research support; BJSM: Publishing royalties, financial or material support; ESSKA: Board or committee member; iBalance: Stock or stock Options; International Society of Arthroscopy, Knee Surgery, and Orthopaedic Sports Medicine: Board or committee member; Journal of Bone and Joint Surgery—American: Editorial or governing board; Knee: Editorial or governing board; Knee Surgery, Sports Traumatology, Arthroscopy: Editorial or governing board; Smith & Nephew: Research support. All other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Full ICMJE author disclosure forms are available for this article online, as supplementary material.

Received February 28, 2023; accepted August 21, 2023.

Address correspondence to Joaquin Sanchez-Sotelo, M.D., Ph.D., Department of Orthopedic Surgery, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, U.S.A. E-mail: sanchezsotelo.joaquin@mayo.edu

0749-8063/23226/$36.00
https://doi.org/10.1016/j.arthro.2023.08.084

As the largest and strongest muscle in the rotator cuff, the subscapularis plays an important role in both stabilization and internal rotation of the glenohumeral joint. Improved understanding in shoulder arthroscopy has led to a focus on identifying subscapularis involvement during arthroscopic rotator cuff repair and techniques to supplement and support rotator cuff repair constructs, such as through comma tissue and its connections with both the subscapularis and supraspinatus.[1-5] Although increased awareness paired with improvements in imaging modalities and the increased adoption of arthroscopic techniques has resulted in a higher rate of subscapularis tear diagnosis and treatment than in the past,[2] the number of "neglected" subscapularis tears remains high, with more than half of all revision rotator cuff repairs having neglected subscapularis tears.[1,4] Furthermore, the retear rate after the repair of neglected subscapularis tears has been shown to be higher than expected, likely because of the tendency of overlooked subscapularis tears to undergo early retraction with progression to muscle atrophy and fatty degeneration.[4,6,7] These findings have highlighted the importance of detecting and treating subscapularis tears through careful preoperative evaluation, as well as during primary arthroscopy.

Despite advances in musculoskeletal imaging, identifying subscapularis tears on preoperative imaging remains challenging, with studies suggesting sensitivities for identifying subscapularis tears on magnetic resonance imaging (MRI) as low as 37%.[8-10] However, recent research has identified a number of indirect but related imaging features on MRI that may be associated with tears of the subscapularis. These include pathologies of the long head of the biceps (LHB), such as dislocation, subluxation, and tearing,[11-14] as well as cystic changes of the lesser tuberosity.[15] Traditionally, clinicians often use more direct signs related to tendon morphology and signal, which include findings such as signal change, abnormal subscapularis tendon length, and the presence and location of fatty atrophy.[16,17] The role of ultrasound scanning in the identification of subscapularis tears has not been well defined and can be operator dependent.[18,19]

In addition to imaging, a variety of clinical tests are commonly used in an attempt to identify and diagnose subscapularis tendon tears preoperatively. These include tests such as the lift-off test, bear-hug test, belly-press test, and internal rotation lag sign. However, like imaging, these physical examination findings have been shown to have relatively low diagnostic sensitivity.[20-22] In a systematic review of clinical tests for diagnosing tears of the subscapularis, pooled sensitivities ranged from 0.32 to 0.55, corresponding to the internal rotation lag sign and bear-hug tests, respectively.[20] The authors concluded that no clinical test is sufficiently reliable for the diagnosis of subscapularis tears before surgery. These findings further highlight the need for better diagnostic methods and tools to identify subscapularis tendon tears before surgery.

In addition to the need for improved identification of subscapularis tears for orthopedic surgeons during preoperative planning and the execution of arthroscopic rotator cuff repairs, there is a need for a tool that can facilitate more accurate tear identification and reporting by radiologists. Particularly in the setting of isolated subscapularis tears, accurate identification of the tear may improve patient care by facilitating earlier referrals and appropriate triaging.

Although traditional statistics may be used to help identify potential risk factors and predictors of intraoperative and postoperative outcomes at a population level, they are constrained by predefined assumptions that make it difficult to apply their findings on an individual, patient-specific level.[23] Machine learning is valuable for its ability to learn from experience and data through pattern recognition and thus generate models that account for nonlinear relationships between multiple patient-specific risk factors.[24-26] As a result, machine learning is capable of identifying those factors that are most important to predict an event (in our study, a subscapularis tear confirmed during surgery) while ignoring those that do not contribute to an accurate prediction or diagnosis.[27] Thus the purpose of this study was to develop a machine learning model capable of identifying subscapularis tears before surgery based on imaging and physical examination findings. We hypothesized that a supervised machine learning model could identify imaging and physical examination findings that would be highly predictive of arthroscopically confirmed subscapularis tendon tears.

Methods

Guidelines

We performed analysis adherent to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines and the Guidelines for Developing and Reporting Machine Learning Models in Biomedical Research.[28,29] The TRIPOD guidelines were developed to promote transparent reporting of prediction model development and validation and include recommendations regarding the description of data sources and participants, proportion of missing data, predictor variables, outcome of interest, statistical analysis, and model development and performance.

Data Source

After institutional review board approval, a prospectively collected database of consecutive patients undergoing arthroscopic rotator cuff repair by a single surgeon between January 2010 and January 2020 was queried. Manual records for this single surgeon series were reviewed for all patients who underwent arthroscopy of the shoulder (1,642 patients). All patients who underwent surgery with this single surgeon and had arthroscopic verification of the condition of the subscapularis in a standardized manner were included. These patients also needed to have preoperative MRI imaging. Patients without preoperative MRI, arthroscopic images, or surgical note documentation of the subscapularis were excluded. After application of these criteria, a total of 202 patients were eligible for inclusion in the current study.

Primary Outcome

The primary outcome of interest was partial or full-thickness tear of the subscapularis tendon confirmed during surgery as verified arthroscopically. Arthroscopic evaluation of subscapularis tears was done in a standardized manner described by Lee et al.[2] We viewed from the posterior portal in beach chair position with both a 30° and 70° arthroscope in 2 main arm positions. We first evaluated the subscapularis in 40° of forward flexion and 30° of internal rotation to visualize the attachment. In addition, we then applied a posterior lever push test in that position along with a posterior humeral translation in 10° of forward flexion and 20° of internal rotation. This was done to ensure that "hidden lesions" were not missed during evaluation, because nearly one fifth of subscapularis tears requiring repair have been shown to be capable of being identified with a 70° arthroscope but not by a 30° arthroscope.[30]

Preoperative Covariates

At the outset, a total of 19 preoperative variables were evaluated for their ability to predict tears of the subscapularis. These were subdivided into imaging-, demographic-, and physical examination-related features. Imaging parameters and the corresponding imaging modality (radiography [XR]; magnetic resonance imaging [MRI]) included the following variables obtained via blinded review of preoperative images by two board-certified radiologists: presence of lesser tuberosity cystic changes (yes or no) (XR), presence of anterior subluxation on axillary view (yes or no) (XR), presence of superior subluxation on Grashey view (yes or no) (XR), presence of abnormal subscapularis tendon length on axial T2 (yes or no) (MRI), presence of lesser tuberosity cystic changes on axial T2 (yes or no) (MRI), medial subluxation or dislocation of the long head of the biceps tendon on axial T2 (none, subluxed, or dislocated) (MRI), type of tear of the long head of the biceps tendon (none, partial, or complete or prior tenotomy) (MRI), presence of subscapularis fatty atrophy on sagittal T1 (yes or no) (MRI), location of subscapularis fatty atrophy (none, upper third, middle third, lower third, full) (MRI), and direct signs related to the presence of a subscapularis tear (changes in tendon morphology and signal) off the lesser tuberosity on sagittal T2 alone (yes or no) (MRI), on axial T2 alone (yes or no) (MRI), or combined (simple diagnosis using direct signs from both views) (yes or no) (MRI). Figure 1 depicts a subset of representative MRI findings associated with tears of the subscapularis tendon that were considered as potential predictive features for the machine learning model. Demographic variables included sex (male or female sex) and age (continuous) at the time of imaging. Physical exam findings included weakness with internal rotation (yes or no), positive lift/push off test (yes or no), positive belly press test (yes or no), positive bear hug test (yes or no), forward flexion angle (continuous), and external rotation angle (continuous). Imaging review was performed by 2 board-certified musculoskeletal radiologists (N.G.R. and C.A.T-H.). Physical examinations were performed by a board-certified shoulder and elbow surgeon (J.S-S.).

Missing Data

To prevent potential biases that may arise from missing data during model development, missing data was treated based on the established method of multiple imputation. Variables with missing data and the proportion with missing data were as follows: presence of weakness with internal rotation (2.0%), lift/push off test (90.6%), belly press test (81.7%), forward flexion angle (3.5%), external rotation angle (2.5%), bear hug test (26.7%), presence of lesser tuberosity cystic change on XR (13.4%), presence of anterior subluxation on axillary XR (12.9%), presence of superior subluxation on Grashey XR (2.0%), dislocation or medial subluxation of long head of biceps on axillary T2 (9.4%), and presence of subscapularis fatty atrophy on sagittal T1 (0.5%). Variables exceeding 30% missing data were removed from further consideration, while variables with less than 30% missing data were imputed and included to determine their predictive value. Stekhoven and Buhlmann[31] established this threshold for missing data by experimentally introducing missing values into training data and comparing imputed values to true values. No significant deviation was found for variables missing less than 30% of values. The "missForrest" multiple imputation method was used to impute missing cases.

Model Development

Following imputation of missing data, feature selection was performed through a combined data-driven and domain expert directed approach. Recursive

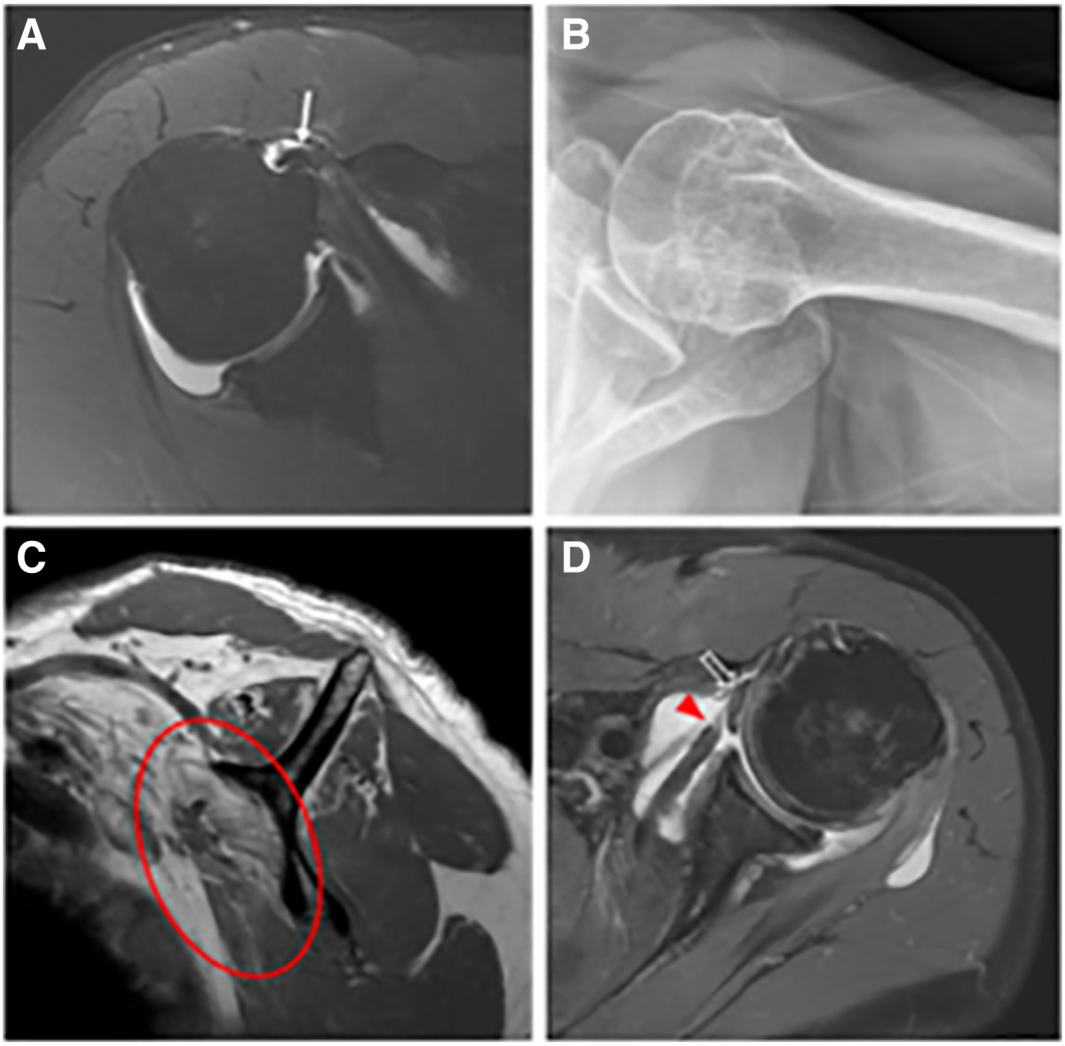

Fig 1. Representative magnetic resonance imaging findings associated with tears of the subscapularis tendon that were considered as potential predictive features for the machine learning model. **(A)** Axial T2 fat-saturated image from a magnetic resonance (MR) arthrogram of the right shoulder that demonstrates subluxation of the biceps tendon into a partial-thickness subscapularis tear (*arrow*). **(B)** Axillary radiograph of the left shoulder demonstrating anterior subluxation of the humeral head relative to the glenoid. **(C)** Sagittal T1-weighted MR image of the left shoulder demonstrating fatty atrophy of the subscapularis muscle belly (*circled*). (D) Axial T2 fat saturated MR image of the left shoulder demonstrating a full-thickness tear of the subscapularis tendon (*arrowhead*), with dislocation of the biceps tendon into the tear (*open arrow*).

feature elimination with random forest algorithms was used to eliminate features with the least predictive value. This method for feature selection begins by first training a model using all features in a dataset, assigning weights and importance scores to each variable, and performing backward selection to remove variables with the lowest importance scores. The model is rebuilt, and this process is repeated in an iterative fashion until the combination of variables that optimizes model accuracy is reached. Selected features were then input into the modeling workflow to train an eXtreme Gradient Boosting (XGBoost) algorithm, which was selected for its predictive power and efficiency. Briefly, the XGBoost algorithm is an ensemble tree-based approach, which sequentially builds trees with the goal of reducing the errors of the previous trees. We used nested cross-validation (CV) to tune hyperparameters and train our model. While this method is computationally expensive, our dataset included relatively few instances, which allowed us to conduct nested cross-validation to prevent biasing the model to the dataset and yielding overly-optimistic results, since non-nested CV uses the same data to tune model parameters and evaluate model performance and can cause data leakage that results in overfitting the data. Simply, outer cross-validation provides the training set for the inner loop, and the inner loop uses this training set to tune

Table 1. Patient Characteristics

Characteristic	Value, Mean (SD) or %
Age	59.1 (10.3)
Sex	
Male	56.9% (115/202)
Female	43.1% (87/202)
Arthroscopic confirmed subscapularis tear	60.0 (121/202)
Upper third	82.6% (100/121)
Upper two thirds	14.0% (17/121)
Full	3.3% (4/121)

SD, standard deviation.

hyperparameters for the model. This tuning process again uses CV on the training set. In the present implementation, 10 folds were used for the outer loop and 4 folds were used for the inner loop. An 80:20 train-test split was used to partition the data into a training set and an independent testing set for internal validation.

Model Evaluation

Model performance was measured using standard methods for assessing machine learning algorithms: discrimination, accuracy, and F1 score (balance between precision and recall). Discrimination measures the diagnostic ability of a binary classifier to correctly label cases of the dependent variables. It is quantified by the c-statistic, which is generated by calculating the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. An ROC AUC of 0.70 to 0.80 is considered acceptable, whereas an ROC AUC of 0.80-0.90 is considered excellent.[32] Accuracy is simply the ratio of true positives and true negatives to the total number of cases evaluated. The F1 score is obtained by calculating the harmonic mean of precision and recall, giving equal weight to each. Using a single score, it allows assessment of both the precision and recall of a machine learning model, where precision is defined as the number of true positives divided by the number of all predicted positive results, and the recall is defined as the number of true-positive results divided by the number of all samples that should have been identified as positive. The F1 score ranges from 0 to 1, with a score of 1 indicating perfect precision and recall and a score of 0 indicating a lack of either precision or recall.

Table 2. Interobserver Agreement of Imaging Findings by κ Values

Parameter	Interobserver Agreement (κ)
XR	
Lesser tuberosity cystic change	0.208
Anterior subluxation on axillary	0.035
Superior subluxation on Grashey	0.304
MRI	
Subscapularis tendon length abnormal on axial T2	0.411
Lesser tuberosity cystic changes on axial T2	0.489
LHB medial subluxation or dislocation on axial T2	0.397
LHB tear on axial T2	0.246
Subscapularis fatty atrophy on sagittal T1	0.465
Location of subscapularis fatty atrophy	0.283
Subscapularis tear on axial T2 (direct signs)	0.562
Subscapularis tear off lesser on sagittal T2 (direct signs)	0.548
Subscapularis tear simple (direct signs)	0.528

LHB, long head of the biceps; MRI, magnetic resonance imaging; XR, x-ray imaging.

Analysis

Continuous variables were assessed with two-sample t-tests, while categorical variables were compared using χ^2 tests. Statistical significance was defined as $P < .05$. Interobserver agreement of radiographic findings was assessed using Fleiss κ statistics. Values of $\kappa < 0.2$ indicate poor agreement; 0.2−0.4, fair agreement; 0.4−0.6, moderate agreement; 0.6−0.8, substantial agreement; and > 0.8, almost perfect agreement. All data analysis was performed with Python 3.8.8.

Results

Study Population Characteristics

We created the cohort through manual identification of patients from a single surgeon's cases, making sure they had available MRI and arthroscopic documentation of the status of the subscapularis. This resulted in a total of 202 shoulders that were included in the machine learning analysis. Of these, 121 (60%) had partial or full thickness tears of the subscapularis verified during surgery. One hundred (82.6%) of these were upper third arthroscopically confirmed subscapularis tears, whereas 17 (14%) were classified as upper two-third tears, and 4 (3.3%) were tears of the whole subscapularis. The mean age of included patients was 59.1 ± 10.3 years. For this cohort, 87 (43.1%) were female (Table 1). Interobserver agreement between 2 board-certified musculoskeletal radiologists ranged from fair to moderate on 11 of 12 imaging parameters, with poor agreement on anterior subluxation as seen on axillary radiographs (Table 2).

Descriptive Analysis

Table 3 provides a descriptive analysis of the study population, including each parameter considered for inclusion in the model, as well as a breakdown of those

Table 3. Descriptive Analysis

Parameter	SST	No SST	All	Missing	*P* Value	OR
MRI quality (3T)	75	56	131	6 (3.0%)	.286	0.716 (0.388-1.323)
Imaging				1 (0.5%)		
XR						
Lesser tuberosity cystic change	15	7	22	1 (0.5%)	.401	1.496 (0.581-3.849)
Anterior subluxation on axillary	15	15	30	1 (0.5%)	.230	0.623 (0.286-1.357)
Superior subluxation on Grashey	18	10	28	1 (0.5%)	.610	1.241 (0.541-2.846)
MRI						
Subscapularis tendon length abnormal on axial T2	36	4	40	1 (0.5%)	<.001	8.153 (2.774-23.961)
Lesser tuberosity cystic changes on axial T2	31	17	48	1 (0.5%)	.448	1.297 (0.662-2.541)
LHB medial subluxation or dislocation on axial T2				1 (0.5%)	<.001	
Normal	49	55	104			
Subluxed	59	24	83			
Dislocated	13	2	15			
LHB tear on axial T2	63	17	80	1 (0.5%)	<.001	4.089 (2.150-7.777)
Subscapularis fatty atrophy on sagittal T1	78	17	95	1 (0.5%)	<.001	6.829 (3.559-13.102)
Location of subscapularis fatty atrophy				1 (0.5%)	<.001	
None	42	64	106			
Upper 1/3	23	5	28			
Upper 2/3	31	7	38			
Full	25	5	30			
Subscapularis tear off lesser on sagittal T2 (direct signs)	90	17	107	1 (0.5%)	<.001	10.930 (5.578-21.418)
Subscapularis tear on axial T2 (direct signs)	102	26	128	1 (0.5%)	<.001	11.356 (5.774-22.334)
Subscapularis tear simple (direct signs)	104	14	118	1 (0.5%)	<.001	29.277 (13.541-63.303)
Demographics						
Age at MRI (yr)	61.3 (9.2)	55.8 (10.9)		0 (0%)	<.001	
Female sex	46	41	87	0 (0%)	.076	0.598 (0.339-1.058)
Physical examination						
FF	141.6° (34.3)	142.3° (35.4)		7 (3.5%)	.450	
ER	58.8° (17.0)	58.1° (21.6)		5 (2.5%)	.406	
IR				10 (5.0%)	.794	
Iliac crest	4	2	6			
SI Joint	22	20	42			
Lumbar	34	15	49			
Thoracic	53	43	96			
Weakness with IR	43	10	53	5 (2.5%)	<.001	4.126 (1.927-8.833)
Lift/Push Off Positive	4	0	4	184 (91.5%)	.018	3.750 (1.620-8.679)
Belly Press Positive	16	3	19	166 (82.6%)	.002	10.667 (2.208-51.533)
Bear Hug Positive	64	26	90	55 (27.4%)	.001	3.249 (1.628-6.486)

ER, external rotation; FF, forward flexion angle; IR, internal rotation; LHB, lateral head of the biceps; MRI, magnetic resonance imaging; SI, sacroiliac; SST, subscapularis tear; XR, x-ray imaging.

patients with and without subscapularis tears who had positive values for each parameter. Preoperative imaging features predictive of an arthroscopically confirmed subscapularis tear include abnormal subscapularis tendon length on axial T2 MRI ($P < .001$; odds ratio [OR] = 8.153; 95% confidence interval [CI] = 2.774-23.961), tear of the long head of the biceps on axial T2 MRI ($P < .001$; OR = 4.089; 95% CI = 2.150-7.777), and subscapularis fatty atrophy on sagittal T1 MRI ($P < .001$; OR = 6.829; 95% CI = 3.559-13.102). Direct signs of tearing including changes in tendon morphology and signal on MRI were also highly predictive of arthroscopically confirmed subscapularis tendon tears ($P < .001$; OR = 29.277; 95% CI = 13.541-63.303). Physical examination parameters predictive of an arthroscopically confirmed subscapularis tear include weakness with internal rotation ($P < .001$; OR = 4.126; 95% CI = 1.927-8.833), a positive lift/push-off test result ($P = .018$; OR = 3.750; 95% CI = 1.620-8.679), a positive belly press test result ($P = .002$; OR = 10.667; 95% CI = 2.208-51.533), and a positive bear hug test result ($P = .001$; OR = 3.249; 95% CI = 1.628-6.486). Older age was also found to be predictive of a subscapularis tear ($P < .001$; mean age (standard deviation) = 61.3 (9.2) years vs. 55.8 (10.9) years).

Preliminary Feature Ranking and Selection

Modeling using exclusively preoperative imaging parameters was performed first. The most important contributors to predicting an arthroscopically confirmed tear of the subscapularis as determined by recursive feature selection with random forest algorithms were (1) subscapularis fatty atrophy on sagittal T1 MRI, (2) tear of the long head of the biceps on axial T2 MRI, (3) abnormal subscapularis tendon length on axial T2 MRI, and (4) lesser tuberosity cystic changes (Fig 2). When features not directly related to the appearance of the rotator cuff and related structures on imaging were added to the model, the most important features for prediction were (1) subscapularis fatty atrophy on sagittal T1 MRI, (2) forward flexion angle, (3) tearing of the long head of the biceps on Axial T2 MRI, (4) external rotation angle, and (5) MRI quality being 3T (Fig 2). Finally, when diagnosis by direct signs visible on MRI—either off the lesser tuberosity on sagittal T2 MRI, on axial T2 MRI, or combined (simple diagnosis using direct signs from both)—were included for possible incorporation in the final model, feature ranking showed that direct changes in tendon morphology and signal were most predictive, because direct signs from both views proved to be the most predictive factor, followed by direct signs on axial T2 alone and direct signs off the lesser tuberosity on sagittal T2 alone.

Algorithm Performance

The final list of variables identified for prediction of arthroscopically confirmed subscapularis tendon tears is shown in Figure 3. XGBoost algorithms provide feature importance scores via information gain, which is defined as the reduction in entropy that results from transforming a dataset in some way. Simply, information gain is a measure of the improvement in accuracy when a decision tree branches on a specific feature. Features shown in Figure 3 are ranked by information gain.

Using only preoperative imaging-related parameters, the XGBoost model demonstrated excellent performance at predicting subscapularis tendon tears (c-statistic, 0.84; accuracy, 0.85; F1 score, 0.87). The model's accuracy did not significantly increase with the addition of patient or physical exam parameters. Further, isolating the model to only the top five features did not substantially alter model performance (c-statistic, 0.84; accuracy, 0.85; F1 score, 0.86). ROC curves demonstrated excellent model concordance and stability and are shown in Figure 4.

Discussion

The main findings of this study were that a machine learning model can predict subscapularis tendon tears in 85% of patients indicated for arthroscopic rotator cuff repair, and that the most important factors for prediction of subscapularis tears involved direct changes in tendon morphology and signal visible on MRI, pathologies related to the long head of the biceps, and the field strength of the MRI itself. An abundance of prior work has shown that subscapularis tendon tears are especially difficult to identify and diagnose prior to arthroscopic evaluation, with both singular imaging and physical examination findings shown to have very low sensitivities.[8-10,20-22] Thus, to aid in preoperative planning and reduce the risk for retear after arthroscopic rotator cuff repair, there is enormous need for a tool to aid clinicians in the identification and diagnosis of subscapularis tendon tears preoperatively. The present analysis shows the feasibility of machine learning to improve diagnoses and outcomes in patients with subscapularis tears undergoing arthroscopic rotator cuff repair.

A number of prior, non-machine learning investigations have studied the accuracy of MRI in predicting subscapularis lesions. Adams et al.[33] evaluated the accuracy of preoperative MRI in predicting a subscapularis tendon tear based on arthroscopy and found that although all 16 patients with preoperative MRI scans that were interpreted by radiologists as positive for subscapularis tendon tears were confirmed to be positive by arthroscopy, the radiologists diagnosed only 16 of 44 (36%) subscapularis tears identified by arthroscopy, resulting in an overall accuracy of 69%.[33] In another investigation, Lenz et al.[8] found that subscapularis tendon tears verified by arthroscopy were correctly identified in only 37 of 97 (38%) cases in the written report of the preoperative MRI.[8] Foad et al.[9] found that subscapularis tendon tears were missed on preoperative MR scanning (MRI and MR arthrogram) in 25 of 40 shoulders (62.5%). Finally, a study with the purpose of determining and proposing a systematic approach to evaluating MRI scans for subscapularis tears found that orthopedic surgeons correctly diagnosed 60 of 82 patients (73%) with subscapularis tendon tears on preoperative MRI that were subsequently identified by arthroscopy.[10] Our machine learning model achieved an accuracy of 85%, which is substantially higher than those reported in prior, non-machine learning studies. By considering nonlinear relationships between predictive features and learning from a relatively large set of data, the machine learning

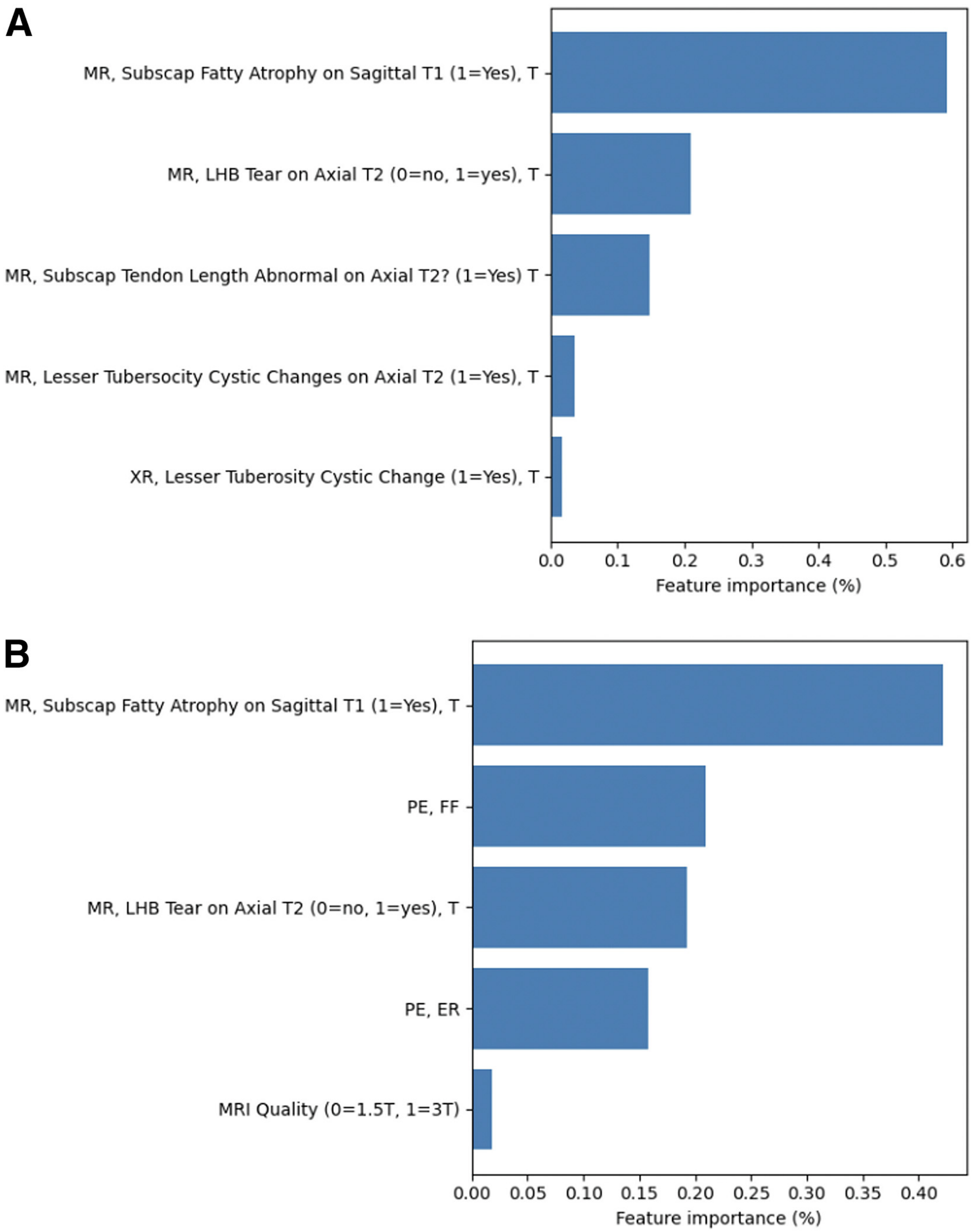

Fig 2. Feature importance as determined using recursive feature elimination with random forest algorithms. **(A)** Feature importance evaluated using exclusively parameters related to the appearance of the rotator cuff and related structures on imaging. **(B)** Feature importance considering all potential variables. ER, external rotation; FF, forward flexion angle; LHB, long head of the biceps; MR, magnetic resonance imaging; PE, physical exam; XR, x-ray imaging.

model developed in this study provides a first step toward improving and assisting humans' ability to diagnose subscapularis tears on preoperative MRI.

In addition to direct signs of tearing commonly used to identify subscapularis tendon tears on MRI, indirect signs such as LHB dislocation, subluxation, and/or tearing and cystic changes of the lesser tuberosity have been increasingly investigated as potential MRI features to aid in the prediction of subscapularis tendon tears.[11-15] Although changes in the morphology and signal of the subscapularis tendon itself as viewed on MRI were found to be the most predictive factors of subscapularis tendon tears, our model showed that LHB pathology does have predictive value for subscapularis

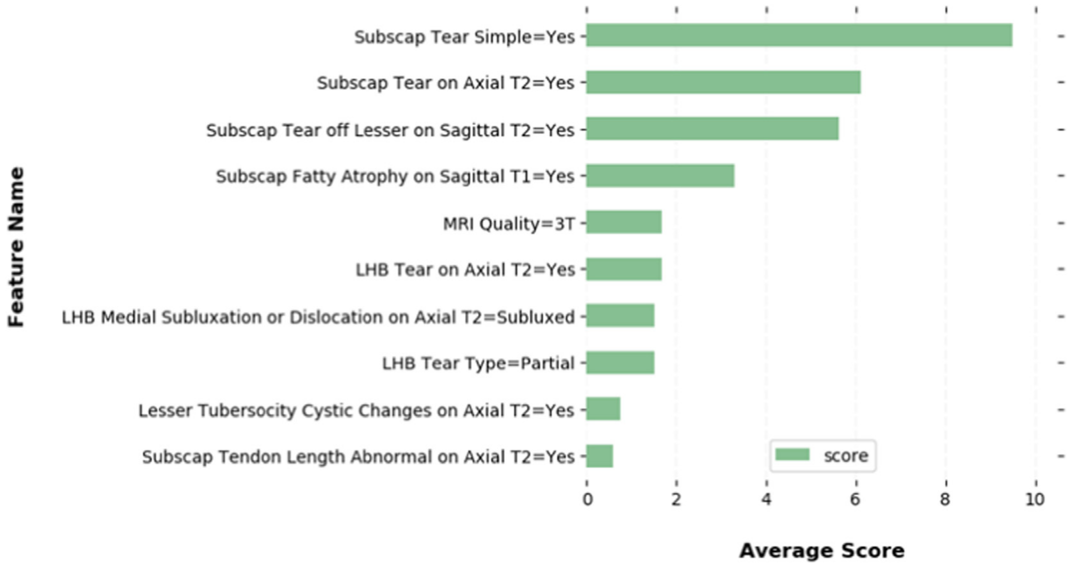

Fig 3. Machine learning model features are ranked by name and average relative importance score. XGBoost algorithms provide feature importance scores via information gain, which measures the improvement in accuracy when a decision tree branches on a specific feature. In this model, direct signs related to the morphology and signal of the tendon on MRI were most predictive, although MRI quality and biceps pathology also showed predictive value for arthroscopically confirmed subscapularis tendon tears. LHB, long head of the biceps; MRI, magnetic resonance imaging.

tendon tears, with visible LHB tearing being one of the top five features required in the final model. This makes sense when considering the anatomy of the rotator interval and biceps pulley system. The superior fibers of the subscapularis along with the anterior fibers of the supraspinatus contribute to the rotator interval. It is in the rotator interval where the superior fibers of the subscapularis tendon converge with the coracohumeral ligament and the superior glenohumeral ligament to form the biceps pulley. Because the subscapularis tendon is an important stabilizer of the LHB, tears of the subscapularis tendon can cause instability and tearing of the LHB.[34]

Interestingly, the field strength of the MRI used to evaluate the patient's rotator cuff was an important predictor of an arthroscopically confirmed tear of the subscapularis. Although the potential reasons for this are multifactorial, they likely involve factors related to the enhanced diagnostic ability of 3T versus 1.5T MRI scanners.[35] Additionally, given the recently increased

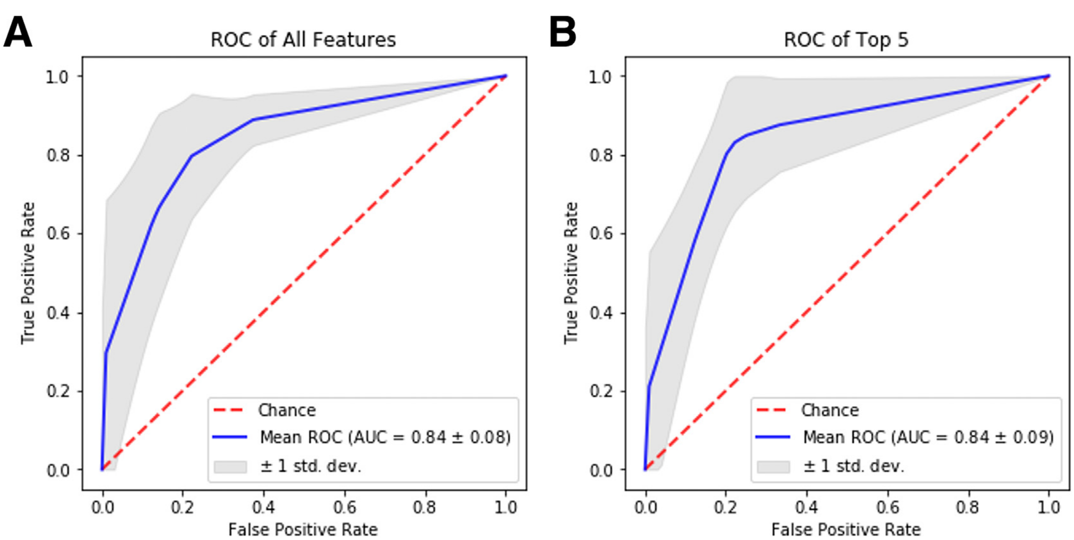

Fig 4. Receiver operating characteristic (ROC) curves of machine learning models created using **(A)** all identified features compared to **(B)** only the top 5 features. The shaded areas represent one standard deviation confidence interval. The area under the curve (AUC) of both models was 0.84.

focus on identifying and diagnosing subscapularis tendon tears before surgery, clinicians with increased suspicion of subscapularis tendon tears based on physical examination findings may be more likely to send "at-risk" patients to imaging centers with higher-resolution scanners that provide a better chance of identifying a torn subscapularis tendon.[35,36]

Interobserver agreement on the presence of specific imaging features related to the tearing of the subscapularis ranged from poor to moderate in our study, further highlighting the difficult task of identifying indicators of subscapularis tendon tears on MRI. The results of this analysis may serve as a proxy for the difficulty of identifying each imaging parameter and be used to help inform surgeons and radiologists of certain imaging findings that may be easier or more difficult to identify than others. For example, the presence of anterior subluxation as viewed on axillary radiography demonstrated the worst interobserver agreement, suggesting that it is highly difficult to objectively classify this parameter on radiography. As a result, it can be concluded that this may not be a reliable or useful sign for predicting subscapularis tears. Not surprisingly, this feature did not show predictive value for our machine learning model. In contrast, direct signs of a torn subscapularis involving changes in morphology and signal of the tendon demonstrated some of the highest rates of interobserver agreement, and thus may serve as reliable indicators of a torn subscapularis tendon. These features were also among the most important predictors of subscapularis tendon tears for our machine learning model.

It is important to note that physical examination findings were not among the most important features for predicting subscapularis tears as determined by the present model. This may be related to their generally low sensitivity for detecting subscapularis tears, as discussed previously. However, given the difficulty of performing some of the clinical tests and the fact that obtaining accurate results from certain tests may be limited by pain, it is not possible to completely rule out their clinical utility in informing a clinician's diagnosis of subscapularis tears preoperatively. Although larger studies with more standardized examinations are needed, findings from high-quality imaging were determined to be more reliable predictors of subscapularis tears than physical exam findings according to the results of the current model. Future modeling with a more complete set of physical examination data is needed to rule out the potential value of the clinical examination when compared to advanced imaging for preoperative predictive models of subscapularis tears.

It is important to note that the model in this study was not trained to distinguish between partial and full-thickness tears. This decision was made to develop a model that could identify features predictive of any tearing of the subscapularis. Although the decision to pursue operative intervention for a subscapularis tendon tear may be based in part on whether the patient experienced a partial or full-thickness tear, there are many other factors that influence whether a patient will benefit from operative intervention for a subscapularis lesion. A model predictive of patients who would benefit from operative intervention would require analysis of patient-specific features, lesion characteristics, as well as post-operative surgical and patient-reported outcomes for a large volume of patients, which was outside the scope of the present study. Surgical treatment is at the discretion of the surgeon, but partial tears that did not respond to nonoperative management (physical therapy, injections, anti-inflammatories) can be considered for surgical management. A goal for future models dedicated to the identification of patients with partial tears who may benefit from operative intervention would be to reduce spent time and costs associated with a trial of nonoperative management. Cost-effectiveness studies could demonstrate the utility of such models. Finally, it is also important to note that this machine learning model was created and internally validated at a single institution. Before widespread adoption in a clinical setting, the model must be tested using a heterogenous population of patients from multiple regions and institutions.

Limitations

There are limitations of this study that warrant further discussion to ensure interpretation of its findings in their proper context. First, as stated above, this model was created using only imaging and physical examination findings to help clinicians identify those features with the most predictive utility for diagnosing subscapularis tears preoperatively. There are, of course, many other potential variables that may provide predictive value for diagnosing subscapularis tendon tears preoperatively. These include patient comorbidities, information on the mechanism of injury, indication for rotator cuff repair, any concomitant shoulder pathologies, prior surgical history, and other demographic variables. In addition, there was a high rate of missing data among certain physical exam tests which required their exclusion from consideration in the machine learning analysis (lift/push off test and belly press test). It is possible that these tests have predictive value for subscapularis tendon tears but were not able to be evaluated in the current study.

Conclusions

In this study, machine learning was successful in predicting subscapularis tears by MRI alone in 85% of patients, and this accuracy did not decrease by isolating the model to the top features. The top 5 features included direct signs related to the presence of tearing as evidenced on MRI (changes in tendon morphology and signal), as well as the quality of the MRI and biceps pathology. Last, in advanced modeling, the addition of physical examination or patient characteristics did not make a significant difference in the predictive ability of this model.

References

1. Lädermann A, Denard PJ, Burkhart SS. Midterm outcome of arthroscopic revision repair of massive and nonmassive rotator cuff tears. *Arthroscopy* 2011;27:1620-1627.
2. Lee J, Shukla DR, Sánchez-Sotelo J. Subscapularis tears: Hidden and forgotten no more. *JSES Open Access* 2018;2: 74-83.
3. Lee SH, Nam DJ, Kim SJ, Kim JW. Comparison of clinical and structural outcomes by subscapularis tendon status in massive rotator cuff tear. *Am J Sports Med* 2017;45: 2555-2562.
4. Yoon TH, Kim SJ, Choi YR, Cho JT, Chun YM. Arthroscopic revision rotator cuff repair: The role of previously neglected subscapularis tears. *Am J Sports Med* 2021;49: 3952-3958.
5. Aiken AJ, Field LD. Techniques and tips for identification of comma tissue in subscapularis tears. *Arthrosc Tech* 2020;9:e859-e862.
6. Bartl C, Salzmann GM, Seppel G, et al. Subscapularis function and structural integrity after arthroscopic repair of isolated subscapularis tears. *Am J Sports Med* 2011;39: 1255-1262.
7. Bartl C, Scheibel M, Magosch P, Lichtenberg S, Habermeyer P. Open repair of isolated traumatic subscapularis tendon tears. *Am J Sports Med* 2011;39:490-496.
8. Lenz R, Kircher J, Schwalba K, Weber MA, Tischer T. Subscapularis tendon tears—Usefulness of written MRI reports for guiding patient referral to shoulder specialists. *Rofo* 2021;193:797-803.
9. Foad A, Wijdicks CA. The accuracy of magnetic resonance imaging and magnetic resonance arthrogram versus arthroscopy in the diagnosis of subscapularis tendon injury. *Arthroscopy* 2012;28:636-641.
10. Adams CR, Brady PC, Koo SS, et al. A systematic approach for diagnosing subscapularis tendon tears with preoperative magnetic resonance imaging scans. *Arthroscopy* 2012;28:1592-1600.
11. Mehta SK, Teefey SA, Middleton W, Steger-May K, Sefko JA, Keener JD. Prevalence and risk factors for development of subscapularis and biceps pathology in shoulders with degenerative rotator cuff disease: A prospective cohort evaluation. *J Shoulder Elbow Surg* 2020;29: 451-458.
12. Yoon JS, Kim SJ, Choi YR, Lee W, Kim SH, Chun YM. Medial subluxation or dislocation of the biceps on magnetic resonance arthrography is reliably correlated with concurrent subscapularis full-thickness tears confirmed arthroscopically. *Biomed Res Int* 2018;2018:2674061.
13. Godenèche A, Nové-Josserand L, Audebert S, et al. Relationship between subscapularis tears and injuries to the biceps pulley. *Knee Surg Sports Traumatol Arthrosc* 2017;25:2114-2120.
14. Lenart BA, Ticker JB. Subscapularis tendon tears: Management and arthroscopic repair. *EFORT Open Rev* 2017;2: 484-495.
15. Cetinkaya M, Öner AY, Ataoglu MB, Ozer M, Ayanoglu T, Kanatli U. Lesser tuberosity cysts and their relationship with subscapularis tears and subcoracoid impingement. *J Orthop Sci* 2017;22:63-68.
16. Ramadan LB, Baptista E, Souza FF, et al. Diagnostic accuracy of preoperative magnetic resonance imaging for detecting subscapularis tendon tears: A diagnostic test study. *Sao Paulo Med J* 2020;138:310-316.
17. Atinga A, Dwyer T, Theodoropoulos JS, Dekirmendjian K, Naraghi AM, White LM. Preoperative magnetic resonance imaging accurately detects the arthroscopic comma sign in subscapularis tears. *Arthroscopy* 2021;37:3062-3069.
18. van der Kraats AM, Peeters NHC, Janssen ERC, Lambers Heerspink FO. Handheld ultrasound does not replace magnetic resonance imaging for diagnosis of rotator cuff tears. *Arthrosc Sports Med Rehabil* 2023;5:e381-e387.
19. Zhu S, Pu D, Li J, et al. Ultrasonography outperforms magnetic resonance imaging in diagnosing partial-thickness subscapularis tear. *Arthroscopy* 2022;38:278-284.
20. Lädermann A, Collin P, Zbinden O, Meynard T, Saffarini M, Chiu JC. Diagnostic accuracy of clinical tests for subscapularis tears: A systematic review and meta-analysis. *Orthop J Sports Med* 2021;9:23259671211042011.
21. Kappe T, Sgroi M, Reichel H, Daexle M. Diagnostic performance of clinical tests for subscapularis tendon tears. *Knee Surg Sports Traumatol Arthrosc* 2018;26:176-181.
22. Yoon JP, Chung SW, Kim SH, Oh JH. Diagnostic value of four clinical tests for the evaluation of subscapularis integrity. *J Shoulder Elbow Surg* 2013;22:1186-1192.
23. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: Making sense of the differences. *Knee Surg Sports Traumatol Arthrosc* 2022;30:753-757.
24. Oeding JF, Williams RJ, Nwachukwu BU, et al. A practical guide to the development and deployment of deep learning models for the orthopedic surgeon: Part I. *Knee Surg Sports Traumatol Arthrosc* 2023;31:382-389.
25. Pruneski JA, Pareek A, Kunze KN, et al. Supervised machine learning and associated algorithms: Applications in orthopedic surgery. *Knee Surg Sports Traumatol Arthrosc* 2023;31:1196-1202.
26. Eckhardt CM, Madjarova SJ, Williams RJ, et al. Unsupervised machine learning methods and emerging applications in healthcare. *Knee Surg Sports Traumatol Arthrosc* 2023;31:376-381.
27. Oeding JF, Lu Y, Pareek A, et al. Understanding risk for early dislocation resulting in reoperation within 90 days of reverse total shoulder arthroplasty: Extreme rare event detection through cost sensitive machine learning. *J Shoulder Elbow Surg* 2023;32:e437-e450.
28. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg* 2015;102:148-158.
29. Luo W, Phung D, Tran T, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *J Med Internet Res* 2016;18:e323.

30. Kim JH, Do WS, Lim JR, Yoon TH, Chun YM. Subscapularis tendon tears hidden by the medial biceps sling can be missed on arthroscopic examination. *Arch Orthop Trauma Surg* 2023;143:3251-3258.
31. Stekhoven D, Buhlmann P. MissForrest-nonparametric missing value for imputation for mixed-type data. *Bioinformatics* 2012;28:112-118.
32. Hosmer D, Lemeshow S, Sturdivant R. *Applied Logistic Regression*, 3rd Ed. New York: Wiley, 2013.
33. Adams CR, Schoolfield JD, Burkhart SS. Accuracy of preoperative magnetic resonance imaging in predicting a subscapularis tendon tear based on arthroscopy. *Arthroscopy* 2010;26:1427-1433.
34. Kim BR, Lee J, Ahn JM, et al. Predicting the clinically significant subscapularis tendon tear: Malposition and tear of the long head of the biceps tendon on shoulder magnetic resonance imaging. *Acta Radiol* 2021;62:1648-1656.
35. Smith TO, Daniell H, Geere JA, Toms AP, Hing CB. The diagnostic accuracy of MRI for the detection of partial- and full-thickness rotator cuff tears in adults. *Magn Reson Imaging* 2012;30:336-346.
36. Malavolta EA, Assunção JH, Gracitelli MEC, Yen TK, Bordalo-Rodrigues M, Ferreira Neto AA. Accuracy of magnetic resonance imaging (MRI) for subscapularis tear: a systematic review and meta-analysis of diagnostic studies. *Arch Orthop Trauma Surg* 2019;139:659-667.