



Article

Approaching Artificial Intelligence in Orthopaedics: Predictive Analytics and Machine Learning to Prognosticate Arthroscopic Rotator Cuff Surgical Outcomes

Anish G. Potty^{1,2,3,†}, Ajish S. R. Potty^{1,†}, Nicola Maffulli^{4,5,6,7}, Lucas A. Blumenschein⁸, Deepak Ganta^{9,*}, R. Justin Mistovich⁸, Mario Fuentes⁹, Patrick J. Denard¹⁰, Paul M. Sethi¹¹, Anup A. Shah¹² and Ashim Gupta^{1,13,14,15,*}

- ¹ South Texas Orthopedic Research Institute (STORI Inc.), Laredo, TX 78045, USA
 - ² The Institute of Musculoskeletal Excellence (TIME Orthopaedics), Laredo, TX 78041, USA
 - ³ School of Osteopathic Medicine, The University of the Incarnate Word, San Antonio, TX 78209, USA
 - ⁴ Department of Musculoskeletal Disorders, School of Medicine and Surgery, University of Salerno, 84084 Fisciano, Italy
 - ⁵ San Giovanni di Dio e Ruggi D'Aragona Hospital "Clinica Ortopedica" Department, Hospital of Salerno, 84124 Salerno, Italy
 - ⁶ Centre for Sports and Exercise Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London E1 4DG, UK
 - ⁷ School of Pharmacy and Bioengineering, Keele University School of Medicine, Stoke on Trent ST5 5BG, UK
 - ⁸ Department of Orthopaedics, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA
 - ⁹ School of Engineering, Texas A&M International University, Laredo, TX 78041, USA
 - ¹⁰ Southern Oregon Orthopedics, Medford, OR 97504, USA
 - ¹¹ Orthopaedic & Neurosurgery Specialists, Greenwich, CT 06905, USA
 - ¹² Kelsey-Seybold Clinic, Houston, TX 77584, USA
 - ¹³ Future Biologics, Lawrenceville, GA 30043, USA
 - ¹⁴ BioIntegrate, Lawrenceville, GA 30043, USA
 - ¹⁵ Regenerative Orthopaedics, Noida 201301, Uttar Pradesh, India
- * Correspondence: deepak.ganta@tamiu.edu (D.G.); ashim6786@gmail.com (A.G.)
† These authors contributed equally to this work.



Citation: Potty, A.G.; Potty, A.S.R.; Maffulli, N.; Blumenschein, L.A.; Ganta, D.; Mistovich, R.J.; Fuentes, M.; Denard, P.J.; Sethi, P.M.; Shah, A.A.; et al. Approaching Artificial Intelligence in Orthopaedics: Predictive Analytics and Machine Learning to Prognosticate Arthroscopic Rotator Cuff Surgical Outcomes. *J. Clin. Med.* **2023**, *12*, 2369. <https://doi.org/10.3390/jcm12062369>

Academic Editor: Wing Hoi Cheung

Received: 10 February 2023

Revised: 9 March 2023

Accepted: 17 March 2023

Published: 19 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Machine learning (ML) has not yet been used to identify factors predictive for post-operative functional outcomes following arthroscopic rotator cuff repair (ARCR). We propose a novel algorithm to predict ARCR outcomes using machine learning. This is a retrospective cohort study from a prospectively collected database. Data were collected from the Surgical Outcome System Global Registry (Arthrex, Naples, FL, USA). Pre-operative and 3-month, 6-month, and 12-month post-operative American Shoulder and Elbow Surgeons (ASES) scores were collected and used to develop a ML model. Pre-operative factors including demography, comorbidities, cuff tear, tissue quality, and fixation implants were fed to the ML model. The algorithm then produced an expected post-operative ASES score for each patient. The ML-produced scores were compared to actual scores using standard test-train machine learning principles. Overall, 631 patients who underwent shoulder arthroscopy from January 2011 to March 2020 met inclusion criteria for final analysis. A substantial number of the test dataset predictions using the XGBoost algorithm were within the minimal clinically important difference (MCID) and substantial clinical benefit (SCB) thresholds: 67% of the 12-month post-operative predictions were within MCID, while 84% were within SCB. Pre-operative ASES score, pre-operative pain score, body mass index (BMI), age, and tendon quality were the most important features in predicting patient recovery as identified using Shapley additive explanations (SHAP). In conclusion, the proposed novel machine learning algorithm can use pre-operative factors to predict post-operative ASES scores accurately. This can further supplement pre-operative counselling, planning, and resource allocation. Level of Evidence: III.

Keywords: machine learning; artificial intelligence; orthopaedics; arthroscopic rotator cuff repair; functional outcomes; predictive modelling; American Shoulder and Elbow Surgeons (ASES) score

1. Introduction

The optimal management of shoulder disorders depends on recognizing the natural history of disability caused by an injury and the anticipated outcomes after operative treatment when this is indicated. Clinical and patient-reported outcome measures are critical to truly understand post-operative function and monitor the progress of an effective care plan [1]. Using machine learning (ML), the goal of this study was to analyse such factors to determine the factors most predictive for successful outcomes. ML is a novel field of study that employs computer algorithms and statistical analysis to determine complex trends and patterns in the data that may not be easily discernible by humans. ML uses data to build empirical/statistical models to describe the behaviour of a system [2].

As such, there is a growing body of literature on machine learning to analyse data and answer clinical questions for both the diagnosis and prognostication of rotator cuff tears [3–5]. Recent reviews have demonstrated the wide range of potential applications, from the analysis of ultrasound to diagnose rotator cuff tears to the characterization of rotator cuff fatty degeneration on CT (computer tomography) scans [3,5]. Conversely, other studies have attempted to develop a clinical prediction tool to forecast the chance of complications versus clinical improvement following repair [4]. However, while exciting, there is much room for improvement regarding the application and accuracy of such ML models [5].

Focusing on patients with operative rotator cuff pathology, we developed a novel algorithm to predict arthroscopic rotator cuff repair (ARCR) outcomes. We examined pre-operative and post-operative American Shoulder and Elbow Surgeons (ASES) score. The ASES score is a widely reported and validated patient-reported outcome measure (PROM) applicable to all patients with shoulder pathologies and independent of their specific diagnosis [6,7]. Through machine learning, we sought to understand whether specific pre-operative characteristics could ultimately predict recovery.

Rotator cuff tears are an ideal musculoskeletal condition to study with machine learning. Many factors influence treatment and outcomes, complicating the ability of surgeons to predict outcomes reliably. Secondly, the burden of disease is substantial: each year, an estimated 250,000 people in the U.S. suffer rotator cuff tears, resulting in 3–4 billion dollars spent annually [8].

As a result, surgeons may be economically “punished” for taking on more complex cases with the potential for poorer outcomes, and such policies may act as a barrier to care for those who need it most.

However, these questions cannot be adequately answered with conventional statistical methods such as simple linear regression, as large volumes of data can be scattered. Machine learning (ML) algorithms can overcome these limitations, improve prediction accuracy, and reduce the margin of error between actual and predicted data [2,9–12]. This approach was recently applied in orthopaedics to optimize the number of questions in the Knee Injury and Osteoarthritis Outcome Score (KOOS) activities of the daily living questionnaire following knee surgery [2].

The goal of this study is to use ML in prospectively collected pre- and post-operative data of patients who underwent ARCR to develop a novel algorithm to predict arthroscopic rotator cuff outcomes. We hypothesized that ML algorithms could be used to predict 3-, 6-, and 12-month post-operative ASES scores for patients who underwent rotator cuff repair.

2. Materials and Methods

2.1. Study Design

We performed a retrospective review of prospectively collected data of patients who underwent ARCR performed by several surgeons between April 2011 and April 2019. The surgical technique was based on the surgeon’s preference. As the data were collected for a multi-centric database, all surgeries were performed with single- or double-row suture bridge construct with a configuration of 3 medial 2 lateral, 2 medial 2 lateral, 1 medial 2 lateral, and 2 medial 1 lateral for the double-row fixation and 3 and 2 anchors for the single

row construct. Data for analysis was extracted from the Surgical Outcome System (SOS) global registry, an international patient-reported outcome database maintained by Arthrex. No institutional review board (IRB) approval was required, as SOS global registry is IRB approved and adheres to Health Insurance Portability and Accountability Act (HIPAA) regulations. All SOS global registry users have access to the shared deidentified data.

We included patients who had fully documented demographic and surgical data. Patients who did not complete pre-operative, 3-, 6-, and 12-month post-operative ASES surveys were excluded. Additional exclusion criteria included patients who underwent revision surgery, those who lacked complete follow-up at the specified time points, and patients with incomplete questionnaires. The dependent (or target) variables for this study were 3-, 6-, and 12-month post-operative ASES scores. Several patient- and surgery-related independent variables were considered for this multivariate analysis to better understand their impact on the target variables. Patient-related factors examined were gender, age, BMI, tobacco use, and past medical history of diabetes. Surgical-related factors included the number of tendons torn, tendon quality, Cofield tear size, retraction stage, tear shape, medial anchor type, number of medial knotless anchors, number of medial suture anchors, lateral anchor type, number of lateral knotless anchors, number of lateral suture anchors, pre-operative visual analogue pain score (VAPS), pre-operative ASES score, and the year of operation.

The binary features in the dataset (gender, tobacco use, and history of diabetes) were encoded as “0” and “1”. The other categorical features in the dataset were converted to numerical values based on domain understanding to improve model predictions. The following encoding was used: tendon quality (poor: 1, fair: 2, good: 3, excellent: 4); Cofield tear size (small (<1 cm): 1, medium (1–3 cm): 2, large (3–5 cm): 3, massive (>5 cm): 4); retraction stage (stage I: 1, stage II: 2, stage III: 3, stage IV: 4); tear shape (L-shaped posterior: 1, L-shaped anterior: 2, U-shaped: 3, avulsion/crescent: 4, massive contracted: 5, longitudinal: 6); anchor type (suture anchor: 1, knotless anchor: 0, tenodesis screw: 0).

2.2. Data Preparation and Model Building

Data processing, analysis, and ML model building were performed using Python 3.7.4 (<http://www.python.org> accessed on 1 December 2022). Python packages such as matplotlib, NumPy, Pandas, and Scikit-learn were used for data wrangling, statistical analysis, visualization, and ML model building [13]. The surgical outcomes model is a multi-target regression problem, as the goal is to predict patient recovery at multiple time points (3, 6, and 12 months) after surgery using pre-operative information. To achieve this, the `multioutputregressor` function from the scikit-learn library in Python was used to fit multiple target variables. Cross-validation (CV) is a de facto standard to estimate model prediction errors and the most popular approach for model selection and hyperparameter tuning [14]. K-fold cross-validation involves partitioning the dataset into k equal-sized subsets (or folds), training the ML model on all but one subsets (i.e., $k-1$ subsets), and then evaluating the model on the held-out subset. This process is then repeated k times with a different subset held out each time.

The data were randomly split into two sets: a training set with 80% data and a test set with the remaining 20% data. The 80/20 data split is a commonly used ML method and was chosen accordingly [15]. The ML model was trained using the training data, and the model's performance was confirmed on the test data. Each feature, prior to model building, was scaled (forcing the mean to 0 and scaling the variance to 1) to help to better interpret the model results. Several machine learning models (linear regression, ridge regression, lasso, support vector regression, k -nearest neighbour, random forest, and XGBoost) were evaluated in this study, and the best model was selected based on 10-fold CV error (Table 1). Although linear regression, ridge regression, and lasso had lower RMSE, XGBoost was chosen as the “best” model for further refinement based on acceptable root-mean-square error (RMSE) and normally distributed errors (less bias in model predictions). Moreover,

linear regression, ridge regression, and lasso models highly weighted “gender” as the most important predictor of post-operative ASES, which medically seems highly unlikely. Therefore, XGBoost was chosen as the “best” model for this work. Hyperparameter tuning for the XGBoost model was performed to identify the “best” model based on minimizing the RMSE through 5-fold cross-validation [16]. This was accomplished by searching over a grid space of select key XGBoost hyperparameters (learning_rate: 0.001, 0.005, 0.01, 0.1; max_depth: 6, 8, 10; n_estimators: 200, 400, 500, 600, 700; min_child_weight: 0.5, 1, 2; colsample_bytree: 0.3, 0.5) using the “GridSearchCV” object in Scikit-learn. The set of XGBoost hyperparameters that resulted in the lowest cross-validation error were as follows: n_estimators = 400, learning_rate = 0.01, max_depth = 6, min_child_weight: 2, and colsample_bytree = 0.5. The performance of the “best” model was then evaluated on the test dataset to gauge its performance on this blind, held-out data.

Table 1. Root-mean-square error (RMSE) and 95% confidence interval (CI) calculated using 10-fold cross-validation to help identify “best” machine learning model.

Algorithm	10-Fold CV RMSE for 3-Month Post-Operative ASES	95% CI
LASSO	15.25	14.36–16.14
Ridge Regression	15.31	14.47–16.16
Linear Regression	15.32	14.47–16.16
XGBoost	15.90	14.80–17.00
Random Forest	16.35	15.40–17.31
K-Nearest Neighbour (KNN)	17.02	16.04–18.01
Support Vector Regression	17.20	16.00–18.40

To better understand model predictions, the SHapley Additive exPlanations (SHAP) method was used to explain global feature importance and individual predictions [17]. SHAP values were first obtained using the TreeExplainer method to explain every prediction of the XGBoost model. The next step involved plotting the explanations using the “summary_plot” method. Explanation of individual predictions was performed using the “force_plot” method within the SHAP library. To summarize, the methods involved are shown in Figure 1 through a simple illustration.

2.3. Data Analysis

Distributions of raw data were analysed for symmetry and skewness. The data were non-symmetric and left-skewed with a skewness of -0.45 , -1.12 , and -1.92 for the 3-, 6-, and 12-month post-operative ASES scores, respectively, meaning that the tail on the left side (or lower ASES scores) of the distribution is considerably drawn out compared to the right tail (or higher ASES scores) of the distribution (Figure 2). This is not surprising since the expectation is that most patients would experience an improvement in ASES scores after surgery. For reference, a symmetric distribution such as the normal distribution (bell curve) has a skewness of 0. Since the target (or independent) variables (i.e., post-operative ASES scores) are left-skewed, the machine learning model is trained on an imbalanced dataset and is less likely to accurately predict the outcome in patients with low post-operative ASES scores. Similarly, if the target variable were right-skewed, the model would be less likely to correctly predict cases with high post-operative scores. To reduce target variable skewness and improve model predictions, a mathematical transformation was performed by subtracting the pre-operative score from the post-operative score. This transformed variable demonstrates the improvement in post-operative ASES scores compared to the pre-operative scores. The post-transformation distributions for 3, 6, and 12 months post operation had a corresponding skewness of 0.06, -0.14 , and -0.24 . In other words, the skewness is closer to zero, indicating that these distributions are more symmetric and closer

to a normal distribution. Other mathematical transformations such as log, square, and square root were not as effective on this dataset.

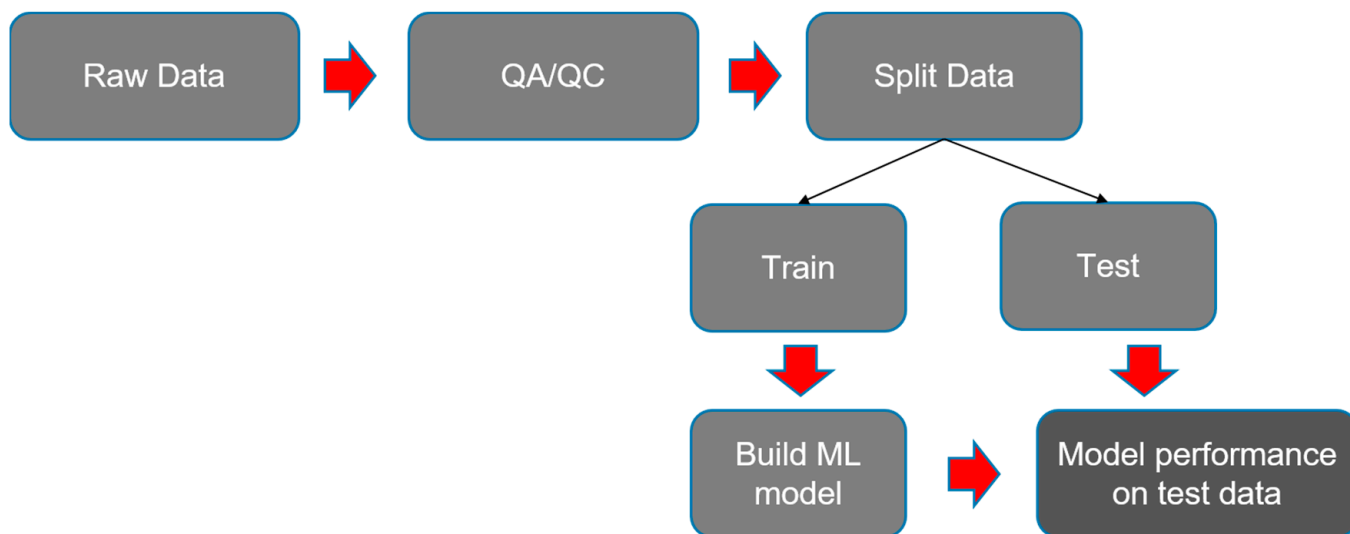


Figure 1. Simple illustration of various steps and their dependence on the machine learning process description.

The model building and model selection steps were undertaken using the transformed target variable described in the Methods section (Section 2.2 data preparation and model building). Root-mean-square error (RMSE) is a measure of the standard deviation of the prediction errors and is a commonly used heuristic to evaluate machine learning models. The average RMSE, calculated using 10-fold CV, for the various ML models (with default hyperparameters) studied in this work was between 15.3–17.2 for the 3-month post-operative ASES score predictions. For the XGBoost model, the RMSE along with 95% confidence interval calculated using 10-fold cross validation of the training data was 15.90 (95% CI: 14.80–17.00), 16.36 (95% CI: 15.70–17.02), and 14.60 (95% CI: 12.84–16.36) for 3, 6, and 12 months post operation, respectively [18]. The RMSE for the test dataset prediction was 16.50, 14.75, and 12.94 for 3, 6, and 12 months post operation, respectively (Table 2).

Table 2. Root-mean-square error (RMSE) calculated using 10-fold cross-validation for train and test datasets at different post-operative timelines.

Post-Operative Time	Train RMSE (10-Fold CV) in ASES	Test RMSE in ASES
3 months	15.90	16.50
6 months	16.36	14.75
12 months	14.60	12.94

The minimal clinically important difference (MCID) and substantial clinical benefit (SCB) values were obtained from the literature and are also shown in Figure 3 [19]. MCID and SCB were used to characterize the extent of error in model predictions. MCID, in this case, is the smallest change in ASES score that a patient would perceive as meaningful. While MCID is defined as the minimum improvement threshold, SCB indicates a substantial change in clinical state as perceived by the patient.

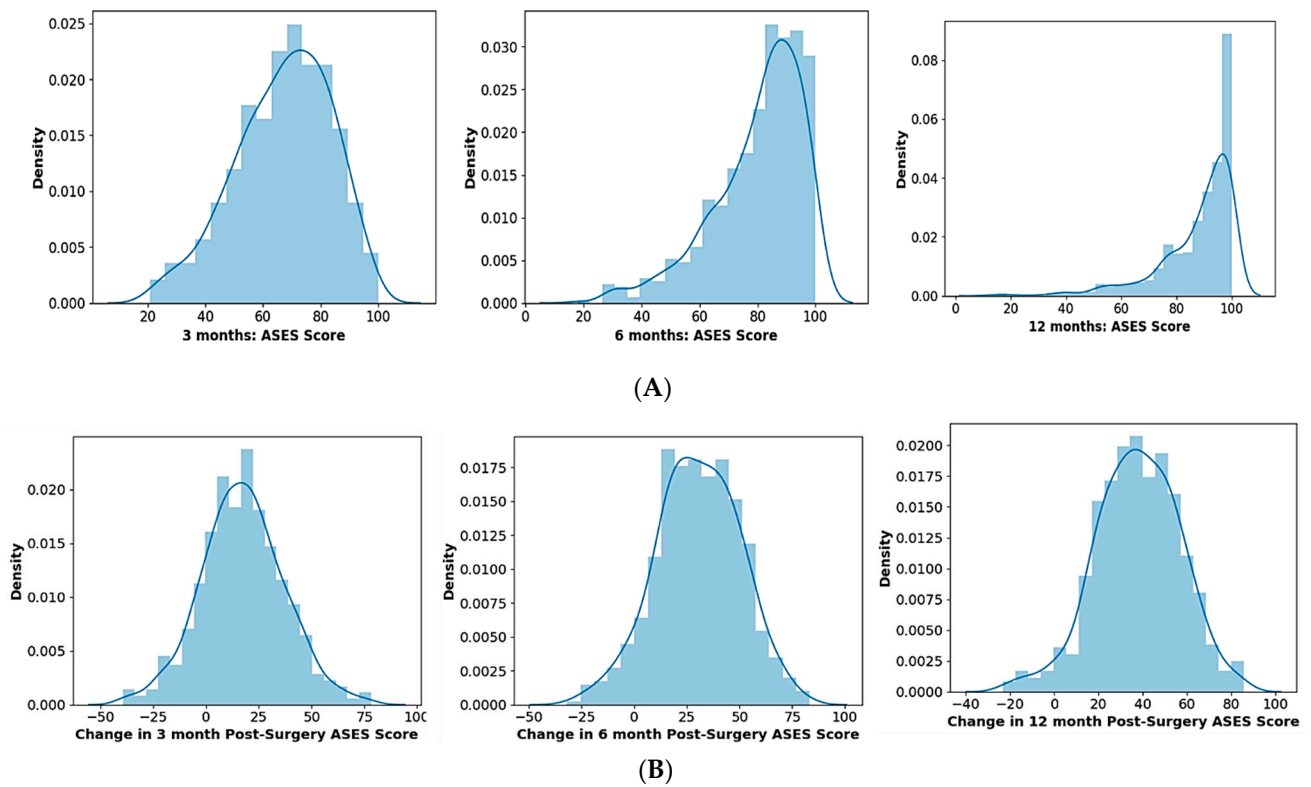


Figure 2. Histogram of (A) 3-, 6-, and 12-month post-operative ASES scores and (B) change in 3-, 6-, and 12-month post-operative ASES scores compared to pre-operative ASES score.

SHAP is a widely used, game-theory-based approach to explain global and local model behaviours [16]. SHAP was used to indicate which features are the most predictive of outcomes. SHAP mean value was used to rank variables or features by highest to least impact on the target variable (i.e., change in ASES score) for the dataset.

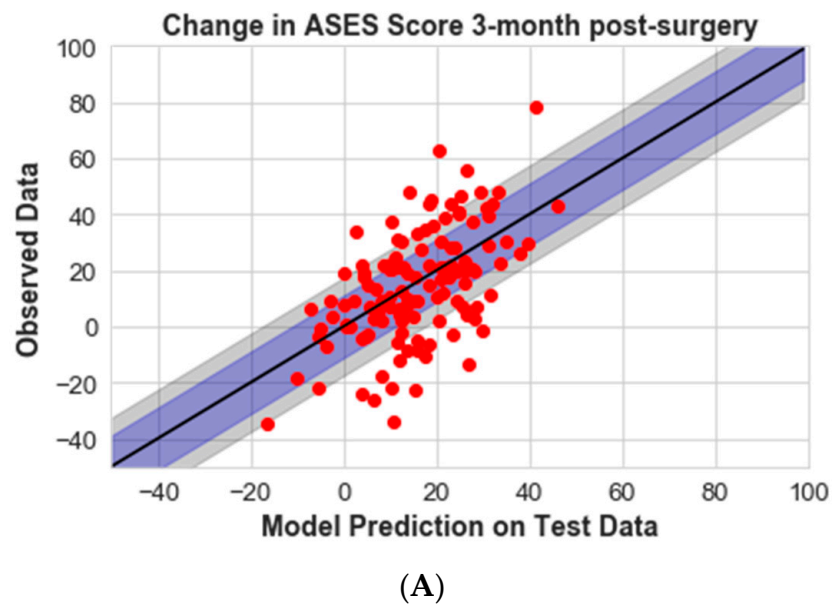


Figure 3. Cont.

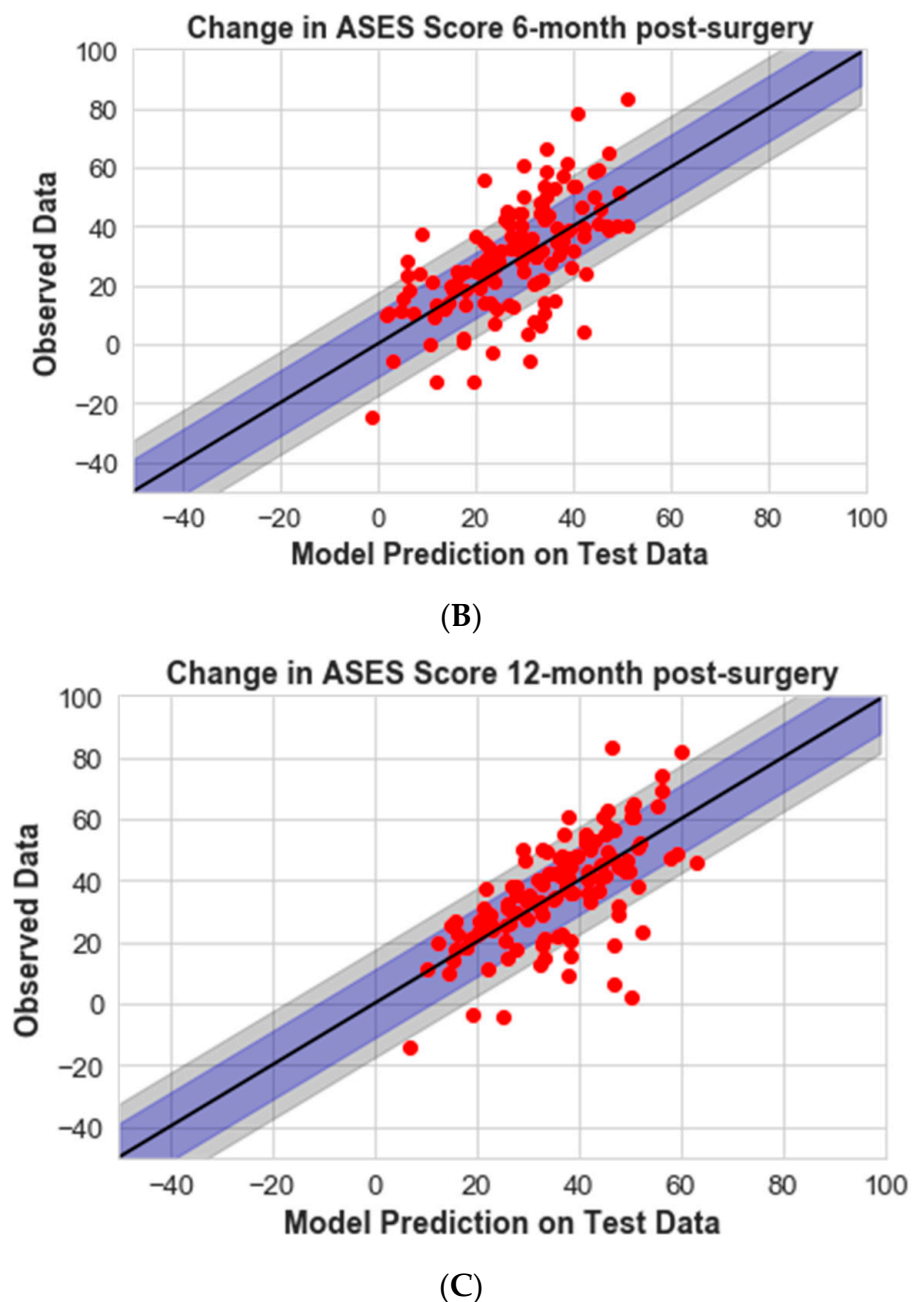


Figure 3. Parity plot of observed and predicted change in ASES score for 3- (A), 6- (B), and 12-month (C) post-operative follow-up relative to the pre-operative ASES score for the test dataset. The shaded blue region is the MCID, and the shaded grey region is the SCB.

3. Results

A total of 4729 patients who underwent ARCR were identified. There were 631 patients between the ages of 24–83 years who met inclusion criteria for analysis (mean age 61.5 years). Of the 631 patients, 362 were males (57%), and 269 were females (43%).

Pre-operative ASES score distribution had a mean and standard deviation of 50 and 17, respectively (Figure 4). The mean post-operative scores increased with recovery time, with the data demonstrating a mean and standard deviation for the 12-month post-operative ASES scores of 87 and 12, respectively, demonstrating a mean improvement in functional outcomes after surgery.

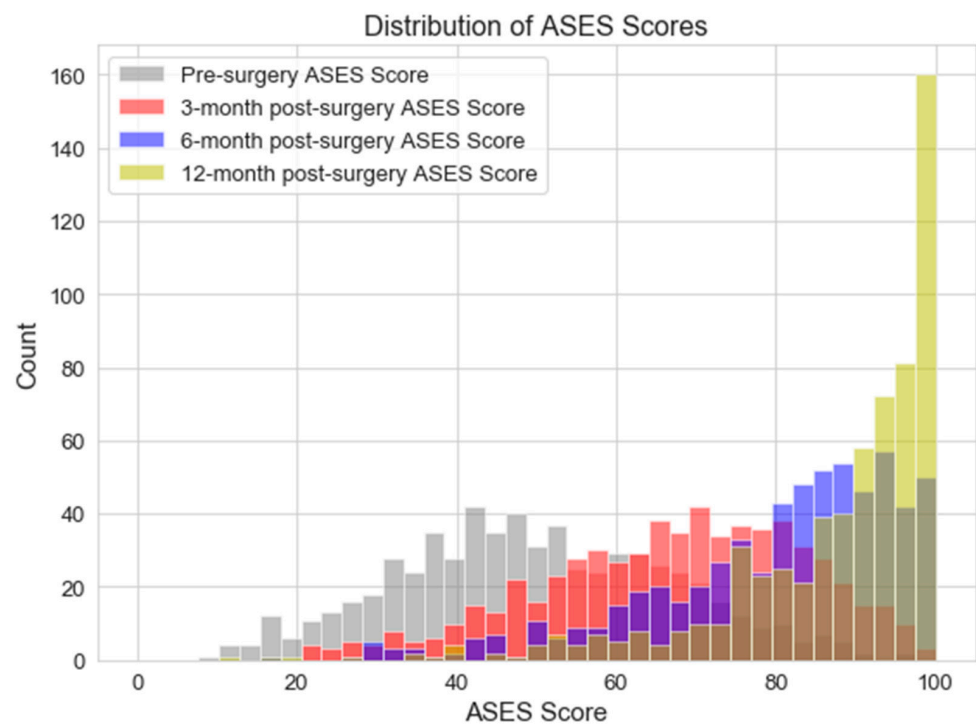


Figure 4. Progressive distribution of the pre-operative, 3-, 6-, and 12-month post-operative ASES scores.

The model demonstrated reasonable performance in predicting recovery progression that healthcare providers could use in their decision-making process. A substantial number of the test dataset predictions using the XGBoost algorithm were within the MCID and SCB thresholds: 69% of the 12-month post-operation predictions were within MCID, while 87% were within SCB, which correlates with the clinical findings of ASES score improvement (Figure 3). The percentage of patients (in the test dataset) predicted within the MCID for 3, 6, and 12 months post operation was 52%, 54%, and 69%, respectively. Similarly, the percentage of patients predicted within SCB for 3, 6, and 12 months post operation was 73%, 78%, and 87%, respectively. The scatter plots comparing model prediction and observed data for 3, 6, and 12 months post operation are demonstrated in Figure 3.

Based on pre-operative information, the model was able to predict 3-, 6-, and 12-month post-operative ASES scores following ARCR (Figure 5). Although the predictions did not exactly match the actuals, the predicted improvement in ASES score lies within the MCID value.

Unlike linear regression models that provide statistical significance (p -value) of each feature as an output, black box ML models (e.g., random forest, XGBoost) output a feature importance score that ranks the relative contribution of each feature towards predicting the target variable. For feature importance, the black box models not only account for main effects but also account for interaction effects between different features. It is important to understand the trade-off between model accuracy and model interpretability. Simple models (e.g., linear regression) have good interpretability but may not have good accuracy. By contrast, complex models usually have better accuracy but poor interpretability (i.e., lack clear explanation of why the model made such a prediction). ML practitioners, having realized the importance of solving this problem, are actively working on various methods to address model interpretability [20]. SHAP is one such approach.

As shown in the SHAP summary values (Figure 6), the features most predictive of post-operative ASES score included the pre-operative ASES score, pre-treatment VAPS, BMI, patient age, and tendon quality. The least predictive factor was patient smoking status (Figure 6). The SHAP summary plot also shows both the positive and negative relationships

of the predictor variables. High pre-operative ASES scores negatively impacted SHAP values. In other words, a patient starting with a high pre-operative ASES score did not have as high of a potential to increase their score further and therefore demonstrated a smaller magnitude of change in the ASES score. In contrast, a patient starting with a low pre-operative ASES score has a larger potential for improvement. Similar conclusions can be drawn for the pre-operative VAPS. Additionally, poor tendon quality leads to slower recovery.

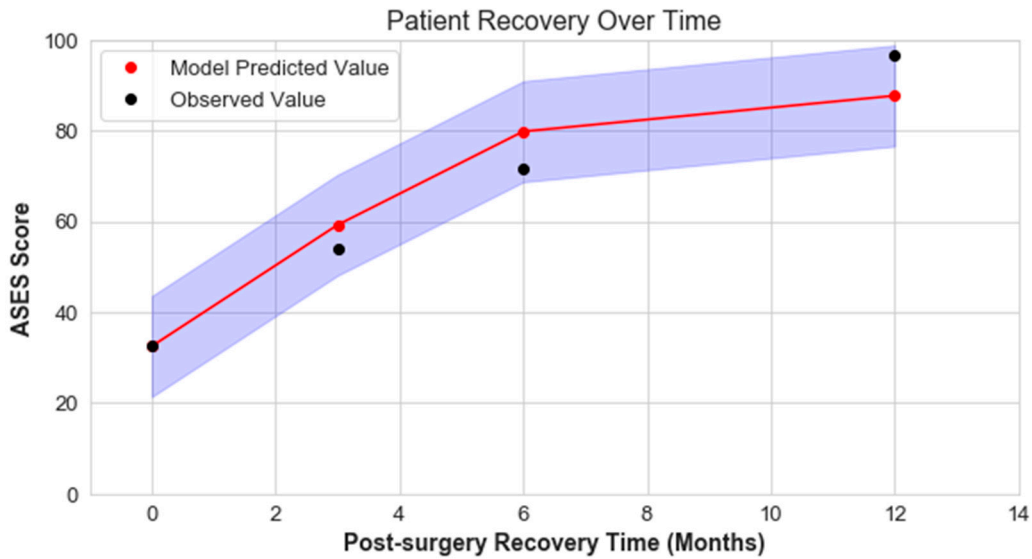


Figure 5. Model predicted 3-, 6-, and 12-month post-operative ASES scores (red) based on pre-operative information for a patient from the test dataset. The observed ASES score (black) is also shown. The shaded region (blue) indicates the bounded area for the MCID value of 11.1.

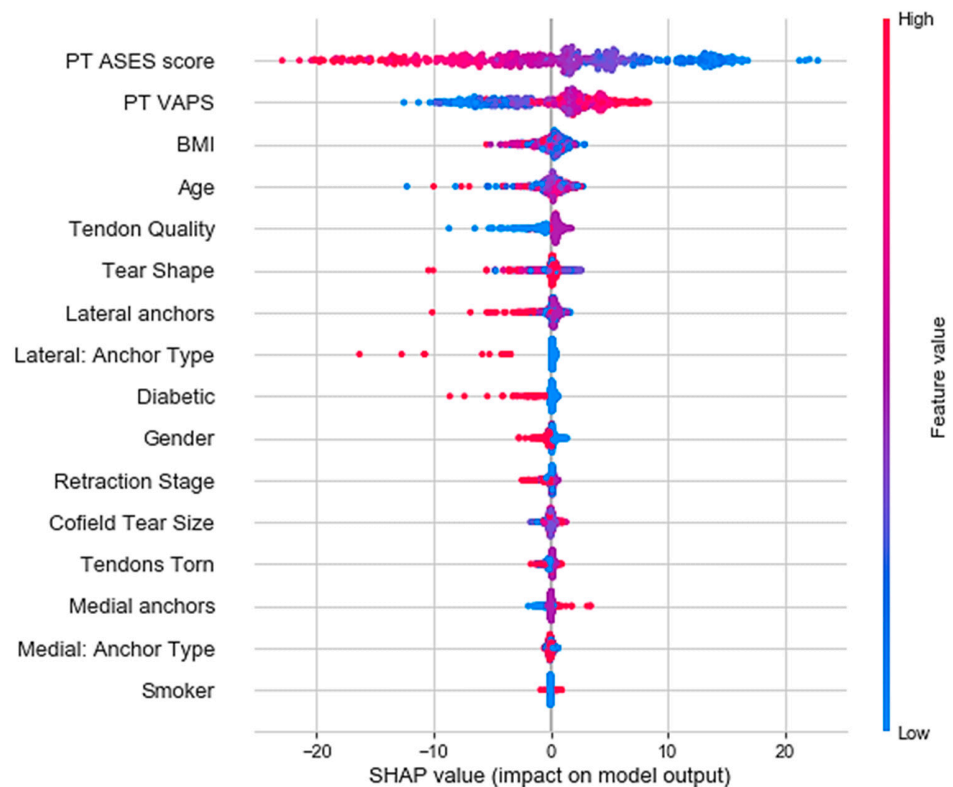


Figure 6. Relative importance of various features on the prediction of 12-month post-operative ASES score.

Case Example

Note that these SHAP summary results are global, i.e., overall trends observed based on this dataset. Interestingly, SHAP can also be used to understand the impact of each factor at an individual patient level. Figure 7 shows the explanation plot generated using the SHAP “force plot” method for one patient in the test dataset. The plot indicates the individual contribution of various factors that lead to the model prediction: improvement in ASES score of 46.99. The base score of 37.32 (the average change in 12-month post-operative ASES score) is also shown in the plot. As shown in Figure 7, the largest contributions in increasing the score come from the pre-operative ASES score and pre-operative VAPS. In addition, the plot attributes having good tendon quality and lower (than average in this dataset) BMI suggest better recovery.

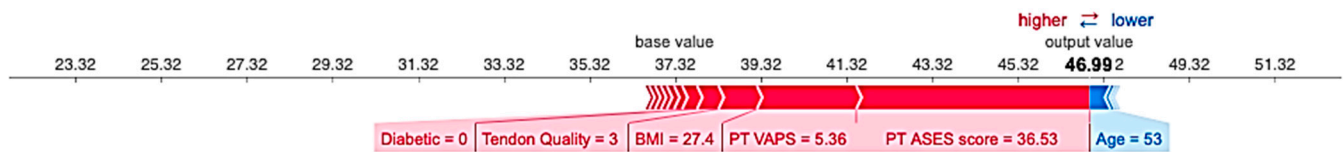


Figure 7. Contribution of various features obtained using SHAP on the prediction of change in 12-month post-operative ASES score for a test dataset patient.

4. Discussion

One of the key findings from this study was that ML could forecast post-operative patient recovery over time using pre-operative factors (Figure 5). We identified pre-operative ASES score, pre-operative VAPS, BMI, age, and tendon quality as key factors impacting patient outcomes (Figure 6). These findings provide a better understanding of the factors influencing surgical outcomes, leading to better informed consent and personalized patient care with data-driven expectations for post-operative recovery.

The application of ML in orthopaedics has recently increased [2,9–12,21]. While the number of orthopaedic studies utilizing ML is still limited, studies from other fields demonstrate its capability to surpass human performance [22–24]. Several orthopaedic studies demonstrate ML can optimize the use of pre-operative assessments and accurately predict the likelihood of patients achieving MCID or SCB in PROMs postoperatively [2,10,21].

One might expect that pre-operative assessment scores would be inversely related to post-operative outcomes; that is, that higher pain and lower functional scores pre-operatively would predict worse post-operative outcomes. Likewise, several studies demonstrate this to be the case [25–28]. Conversely, we found lower pre-operative ASES and VAP scores to be the most significant predictors of post-operative ASES score improvement, with higher pre-operative values negatively influencing post-operative scores. This finding is consistent with Jenssen et al. [29], who analysed factors predictive of post-operative functional outcomes following arthroscopic rotator cuff surgery. They found pre-operative pain scores to be negatively associated with post-operative shoulder function following shoulder arthroplasty. This could be attributed to the fact that patients with worse pre-operative assessments have the greatest potential for improvement. It could also be that patients with worse pre-operative pain can better appreciate their improvements, which reflects the more improved post-operative assessments.

In the present study, individuals with a lower (than average in this dataset) BMI demonstrated better recovery. However, no significant association was found between diabetes mellitus and outcomes. The effects of high BMI, diabetes mellitus, and dyslipidaemia have been previously studied with varied findings. Several studies found that these variables predict worse clinical outcomes, recovery, and tendon healing following rotator cuff injury [30–33]. Conversely, other studies demonstrated no association of BMI with post-operative outcome scores following ARCR [29,34].

Younger age at the time of repair was positively associated with improved outcomes, though the prior literature on this topic is conflicting. Generally, younger age correlates

with more successful recovery following rotator cuff repair [35,36]. Likewise, older age is negatively associated with successful tendon healing, longer recovery time, and increased risk of re-tear following ARCR [37–39]. Other studies demonstrate that when accounting for fatty infiltration, bone mineral density, or retraction of the rotator cuff tendon, there is no independent association between age and rotator cuff healing [40–42]. Furthermore, in other studies, younger age is associated with worse pain and functional outcomes [27,29,43].

In the present investigation, lower pre-operative tendon quality correlated with lower post-operative ASES scores. However, tendon tear size, the number of tendons torn, tear shape, and retraction stage was not strongly associated with a poor outcome. The literature also demonstrates that several measurements of pre-operative tendon quality are associated with worse post-operative outcomes [36]. A larger pre-operative rotator cuff tear size negatively impacts healing, recovery time, functional outcomes, and rate of re-tear [8,32,38,44–46]. Additionally, tendon retraction and fatty infiltration demonstrate a negative impact on healing [34]. Likewise, studies demonstrate that patients with multiple tendon injuries are more likely to develop a rotator cuff defect [47].

Our results did not demonstrate a strong association between gender and post-operative ASES scores, another area of conflicting findings in the literature. Several studies show female sex to be associated with worse post-operative quality of life, mental health, pain, and functional assessments [26,27,48–51]. Conversely, other studies demonstrate that gender does not influence post-operative outcomes [39].

We also did not identify a strong association between tobacco use and post-operative ASES scores. This finding correlates to a prior study demonstrating that tobacco use is not associated with post-operative structural failure following rotator cuff repair [52]. Nonetheless, other studies found that smoking is associated with an increased risk of rotator cuff tears and tear size and worse post-operative clinical outcomes [29,53]. Tobacco use is a modifiable risk factor, and until more definitive data are produced, it is feasible to recommend cessation for the purposes of undergoing ARCR.

5. Limitations

Our results are encouraging, but we are aware of the limitations of the present study. While our sample size is relatively large, there still could be differences between predicted and actual results when applying the model to a real-world population. The model performance (i.e., RMSE) could be further improved by including additional input variables (e.g., compliance with physical therapy protocols and recovery exercises, pain management, and psychological factors) that could impact patient recovery. Additionally, the predicted values from the model may also not represent various regions around the world with different demographics, gender, or ethnic groups. Other features such as patient behavioural factors, medical risk factors, and chronic medical conditions before surgery should also be considered to more accurately predict patient recovery. Additionally, there could also be a recall bias when patients are reporting their outcome questionnaire. However, with more input data, new variables, and model-building refinement, this approach can significantly help surgeons customize their care plans.

One potential application of the model is for surgeons to share the predicted recovery profile with their patients before surgery, thereby setting baseline expectations (Figure 7). Furthermore, the surgeon may run different scenarios—manually or algorithmically—to identify the best course of action for each patient along with any modifiable risk factors. This will allow surgeons to better modulate treatment and rehabilitation techniques with greater confidence.

6. Conclusions

This proposed novel ML algorithm can predict the post-operative ASES scores after ARCR with satisfactory accuracy. While not intended for use in isolation, this model can be used as a critical tool for physicians to formulate better decisions and provide customized, evidence-based care for every patient. In addition, the model may be able to identify

high-risk patients early on and enables surgeons and caregivers to give additional focus to such patients.

Author Contributions: Conceptualization, A.G.P., A.S.R.P. and A.G.; methodology, A.S.R.P.; software, A.S.R.P.; formal analysis, A.G.P., A.S.R.P., D.G., M.F. and A.G.; investigation, A.S.R.P. and M.F.; data curation, A.S.R.P., M.F. and D.G.; writing—original draft preparation, A.G.P., A.S.R.P., L.A.B., M.F. and A.G.; writing—review and editing, A.G.P., A.S.R.P., N.M., D.G., R.J.M., P.J.D., P.M.S., A.A.S. and A.G.; supervision, A.G.; project administration, D.G. and A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study as SOS global registry is IRB-approved and adheres to Health Insurance Portability and Accountability Act (HIPAA) regulations. All SOS global registry users have access to the shared deidentified data.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are contained within the manuscript.

Conflicts of Interest: A.G. has industrial affiliations; however, there are no conflicts of interest with the work presented in this manuscript. Thus, A.G. declares no conflicts of interest. Other authors declare no conflict of interest.

References

1. Michener, L.A.; Valier, A.R.S.; McClure, P.W. Defining substantial clinical benefit for patient-rated outcome tools for shoulder impingement syndrome. *Arch. Phys. Med. Rehab.* **2013**, *94*, 725–730. [[CrossRef](#)] [[PubMed](#)]
2. Gupta, A.; Potty, A.S.; Ganta, D.; Mistovich, R.J.; Penna, S.; Cady, C.; Potty, A.G. Streamlining the KOOS Activities of Daily Living Subscale Using Machine Learning. *Orthop. J. Sports Med.* **2020**, *8*, 2325967120910447. [[CrossRef](#)]
3. Familiari, F.; Galasso, O.; Massazza, F.; Mercurio, M.; Fox, H.; Srikumaran, U.; Gasparini, G. Artificial Intelligence in the Management of Rotator Cuff Tears. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16779. [[CrossRef](#)]
4. Allaart, L.J.H.; Van Spanning, S.; Lafosse, L.; Lafosse, T.; Ladermann, A.; Athwal, G.S.; Hendrickx, L.A.M.; Doornberg, J.N.; Van Den Bekerom, M.P.J.; Buijze, G.A. Developing a machine learning algorithm to predict probability of retear and functional outcomes in patients undergoing rotator cuff repair surgery: Protocol for a retrospective, multicentre study. *BMJ Open* **2023**, *13*, e063673. [[CrossRef](#)]
5. Gupta, P.; Haerberle, H.S.; Zimmer, Z.R.; Levine, W.N.; Williams, R.J.; Ramkumar, P.N. Artificial Intelligence-Based Applications in Shoulder Surgery Leaves Much to Be Desired: A Systematic Review. *JSES Rev. Rep. Tech.* **2023**, *7*, 158–161. [[CrossRef](#)]
6. Michener, L.A.; McClure, P.W.; Sennett, B.J. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: Reliability, validity, and responsiveness. *J. Shoulder Elb. Surg.* **2002**, *11*, 587–594. [[CrossRef](#)]
7. Tashjian, R.Z.; Deloach, J.; Green, A.; Porucznik, C.A.; Powell, A.P. Minimal clinically important differences in ASES and simple shoulder test scores after nonoperative treatment of rotator cuff disease. *J. Bone Jt. Surg.* **2010**, *92*, 296–303. [[CrossRef](#)]
8. Li, L.; Bokshan, S.L.; Ready, L.V.; Owens, B.D. The primary cost drivers of arthroscopic rotator cuff repair surgery: A cost-minimization analysis of 40,618 cases. *J. Shoulder Elb. Surg.* **2019**, *28*, 1977–1982. [[CrossRef](#)] [[PubMed](#)]
9. Cabitza, F.; Locoro, A.; Banfi, G. Machine learning in orthopedics: A literature review. *Front. Bioeng. Biotech.* **2018**, *6*, 75. [[CrossRef](#)]
10. Kumar, V.; Roche, C.; Overman, S.; Simovitch, R.; Flurin, P.H.; Wright, T.; Zuckerman, J.; Routman, H.; Teredesai, A. What Is the Accuracy of Three Different Machine Learning Techniques to Predict Clinical Outcomes after Shoulder Arthroplasty? *Clin. Orthop. Relat. Res.* **2020**, *478*, 2351–2363. [[CrossRef](#)] [[PubMed](#)]
11. Maffulli, N.; Rodriguez, H.C.; Stone, I.W.; Nam, A.; Song, A.; Gupta, M.; Alvarado, R.; Ramon, D.; Gupta, A. Artificial intelligence and machine learning in orthopedic surgery: A systematic review protocol. *J. Orthop. Surg. Res.* **2020**, *15*, 478. [[CrossRef](#)]
12. Kakavas, G.; Malliaropoulos, N.; Pruna, R.; Maffulli, N. Artificial intelligence: A tool for sports trauma prediction. *Injury* **2020**, *51* (Suppl. S3), S63–S65. [[CrossRef](#)]
13. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
14. Bayle, P.; Bayle, A.; Janson, L.; Mackey, L. Cross-Validation Confidence Intervals for Test Error. In Proceedings of the 34th Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–12 December 2020.
15. Dobbin, K.K.; Simon, R.M. Optimally splitting cases for training and testing high dimensional classifiers. *BMC Genom.* **2011**, *4*, 31. [[CrossRef](#)]
16. Putatunda, S.; Rama, K. A Comparative Analysis of Hyperopt as against Other Approaches for Hyper-Parameter Optimization of XGBoost. In Proceedings of the 2018 International Conference on Signal Processing and Machine Learning, ACM, Shanghai China, 28 November 2018; pp. 6–10.

17. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.
18. Pirracchio, R.; Petersen, M.L.; Carone, M.; Rigon, M.R.; Chevret, S.; van der Laan, M.J. Mortality Prediction in Intensive Care Units with the Super ICU Learner Algorithm (SICULA): A Population Based Study. *Lancet Respir. Med.* **2015**, *3*, 42–52. [[CrossRef](#)] [[PubMed](#)]
19. Cvetanovich, G.L.; Gowd, A.K.; Liu, J.N.; Nwachukwu, B.U.; Carbarcas, B.C.; Cole, B.J.; Forsythe, B.; Romeo, A.A.; Verma, N.N. Establishing clinically significant outcome after arthroscopic rotator cuff repair. *J. Shoulder Elb. Surg.* **2019**, *28*, 939–948. [[CrossRef](#)]
20. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2020**, *23*, 18. [[CrossRef](#)] [[PubMed](#)]
21. Fontana, M.A.; Lyman, S.; Sarker, G.K.; Padgett, D.E.; MacLean, C.H. Can Machine Learning Algorithms Predict Which Patients Will Achieve Minimally Clinically Important Differences from Total Joint Arthroplasty? *Clin. Orthop. Relat. Res.* **2019**, *477*, 1267–1279. [[CrossRef](#)]
22. Brynjolfsson, E.; Mitchell, T. What Can Machine Learning Do? Workforce Implications. *Science* **2017**, *358*, 1530–1534. [[CrossRef](#)]
23. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist Level Classification of Skin Cancer with Deep Neural Networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)] [[PubMed](#)]
24. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **2016**, *316*, 2402. [[CrossRef](#)] [[PubMed](#)]
25. Friedman, R.J.; Eichinger, J.; Schoch, B.; Wright, T.; Zuckerman, J.; Flurin, P.-H.; Bolch, C.; Roche, C. Preoperative Parameters that Predict Postoperative Patient-Reported Outcome Measures and Range of Motion with Anatomic and Reverse Total Shoulder Arthroplasty. *J. Shoulder Elb. Surg.* **2019**, *3*, 266–272. [[CrossRef](#)] [[PubMed](#)]
26. Moorthy, V.; Lee, M.; Ang, B.F.H.; Chen, J.Y.; Lie, D.T.T. Comparing the Predictors of Functional Outcomes after Arthroscopic Rotator Cuff Repair Modified Frailty Index, Clinical Frailty Scale, and Charlson Comorbidity Index. *Orthop. J. Sport. Med.* **2021**, *9*, 23259671211005091. [[CrossRef](#)] [[PubMed](#)]
27. Rizvi, S.M.T.; Bishop, M.; Lam, P.H.; Murrell, G.A.C. Factors Predicting Frequency and Severity of Postoperative Pain after Arthroscopic Rotator Cuff Repair Surgery. *Am. J. Sport. Med.* **2021**, *49*, 146–153. [[CrossRef](#)]
28. Wong, S.E.; Zhang, A.L.; Berliner, J.L.; Ma, C.B.; Feeley, B.T. Preoperative Patient-Reported Scores Can Predict Postoperative Outcomes after Shoulder Arthroplasty. *J. Shoulder Elb. Surg.* **2016**, *25*, 913–919. [[CrossRef](#)]
29. Jensen, K.K.; Lundgreen, K.; Madsen, J.E.; Kvakestad, R.; Dimmen, S. Prognostic Factors for Functional Outcome after Rotator Cuff Repair: A Prospective Cohort Study with 2-Year Follow-Up. *Am. J. Sport. Med.* **2018**, *46*, 3463–3470. [[CrossRef](#)]
30. Dhar, Y.; Anakwenze, O.A.; Steele, B.; Lozano Calderon, S.A.; Abboud, J.A. Arthroscopic Rotator Cuff Repair: Impact of Diabetes Mellitus on Patient Outcomes. *Physician Sport.* **2013**, *41*, 22–29. [[CrossRef](#)]
31. Gambhir, N.; Shankar, D.; Alben, M.; Kwon, Y.; Rokito, A.; Virk, M.S. The Effects of Obesity on 1-Year Functional Outcomes after Arthroscopic Rotator Cuff Tear Repair. *J. Shoulder Elb. Surg. Int.* **2022**, *6*, 631–637. [[CrossRef](#)]
32. Kim, Y.-K.; Jung, K.-H.; Kim, J.-W.; Kim, U.-S.; Hwang, D.-H. Factors Affecting Rotator Cuff Integrity after Arthroscopic Repair for Medium-Sized or Larger Cuff Tears: A Retrospective Cohort Study. *J. Shoulder Elb. Surg.* **2018**, *27*, 1012–1020. [[CrossRef](#)]
33. Warrender, W.J.; Brown, O.L.; Abboud, J.A. Outcomes of Arthroscopic Rotator Cuff Repairs in Obese Patients. *J. Shoulder Elb. Surg.* **2011**, *20*, 961–967. [[CrossRef](#)]
34. Fermont, A.J.; Wolterbeek, N.; Wessel, R.N.; Baeyens, J.-P.; de Bie, R.A. Prognostic Factors for Recovery after Arthroscopic Rotator Cuff Repair: A Prognostic Study. *J. Shoulder Elb. Surg.* **2015**, *24*, 1249–1256. [[CrossRef](#)] [[PubMed](#)]
35. Abtahi, A.M. Factors Affecting Healing after Arthroscopic Rotator Cuff Repair. *WJO* **2015**, *6*, 211. [[CrossRef](#)] [[PubMed](#)]
36. Fermont, A.J.M.; Wolterbeek, N.; Wessel, R.N.; Baeyens, J.-P.; de Bie, R.A. Prognostic Factors for Successful Recovery after Arthroscopic Rotator Cuff Repair: A Systematic Literature Review. *J. Orthop. Sport. Phys. Ther.* **2014**, *44*, 153–163. [[CrossRef](#)] [[PubMed](#)]
37. Boileau, P.; Brassart, N.; Watkinson, D.J.; Carles, M.; Hatzidakis, A.M.; Krishnan, S.G. Arthroscopic repair of full thickness tears of the supraspinatus: Does the tendon really heal? *J. Bone Jt. Surg. Am.* **2005**, *87*, 1229–1240. [[CrossRef](#)]
38. Guo, A.A.; Stitz, D.J.; Lam, P.; Murrell, G.A.C. Tear Size and Stiffness Are Important Predictors of Retear: An Assessment of Factors Associated with Repair Integrity at 6 Months in 1526 Rotator Cuff Repairs. *J. Bone Jt. Surg.* **2022**, *7*, e22.00006. [[CrossRef](#)]
39. Manaka, T.; Ito, Y.; Matsumoto, I.; Takaoka, K.; Nakamura, H. Functional Recovery Period after Arthroscopic Rotator Cuff Repair: Is It Predictable Before Surgery? *Clin. Orthop. Relat. Res.* **2011**, *469*, 1660–1666. [[CrossRef](#)]
40. Chung, S.W.; Kim, J.Y.; Kim, M.H.; Kim, S.H.; Oh, J.H. Arthroscopic Repair of Massive Rotator Cuff Tears: Outcome and Analysis of Factors Associated with Healing Failure or Poor Postoperative Function. *Am. J. Sport. Med.* **2013**, *41*, 1674–1683. [[CrossRef](#)]
41. Chung, S.W.; Oh, J.H.; Gong, H.S.; Kim, J.Y.; Kim, S.H. Factors Affecting Rotator Cuff Healing after Arthroscopic Repair: Osteoporosis as One of the Independent Risk Factors. *Am. J. Sport. Med.* **2011**, *39*, 2099–2107. [[CrossRef](#)]
42. Han Oh, J.; Hoon Kim, S.; Kang, J.Y.; Hee Oh, C.; Gong, H.S. Effect of Age on Functional and Structural Outcome after Rotator Cuff Repair. *Am. J. Sport. Med.* **2010**, *38*, 672–678. [[CrossRef](#)]
43. Desai, V.N.; Cheung, E.V. Postoperative Pain Associated with Orthopedic Shoulder and Elbow Surgery: A Prospective Study. *J. Shoulder Elb. Surg.* **2012**, *21*, 441–450. [[CrossRef](#)]

44. Bishop, J.; Klepps, S.; Lo, I.K.; Bird, J.; Gladstone, J.N.; Flatow, E.L. Cuff Integrity after Arthroscopic versus Open Rotator Cuff Repair: A Prospective Study. *J. Shoulder Elb. Surg.* **2006**, *15*, 290–299. [[CrossRef](#)] [[PubMed](#)]
45. Cho, N.S.; Rhee, Y.G. The Factors Affecting the Clinical Outcome and Integrity of Arthroscopically Repaired Rotator Cuff Tears of the Shoulder. *Clin. Orthop. Surg.* **2009**, *1*, 96. [[CrossRef](#)] [[PubMed](#)]
46. Fuchs, B.; Gilbert, M.K.; Hodler, J.; Gerber, C. Clinical and Structural Results of Open Repair of an Isolated One-Tendon Tear of the Rotator Cuff. *J. Bone Jt. Surg.* **2006**, *88*, 309–316. [[CrossRef](#)] [[PubMed](#)]
47. Gulotta, L.V.; Nho, S.J.; Dodson, C.C.; Adler, R.S.; Altchek, D.W.; MacGillivray, J.D. Prospective Evaluation of Arthroscopic Rotator Cuff Repairs at 5 Years: Part II—Prognostic Factors for Clinical and Radiographic Outcomes. *J. Shoulder Elb. Surg.* **2011**, *20*, 941–946. [[CrossRef](#)] [[PubMed](#)]
48. Audigé, L.; Aghlmandi, S.; Grobet, C.; Stojanov, T.; Müller, A.M.; Felsch, Q.; Gleich, J.; Flury, M.; Scheibel, M. Prediction of Shoulder Stiffness after Arthroscopic Rotator Cuff Repair. *Am. J. Sport. Med.* **2021**, *49*, 3030–3039. [[CrossRef](#)]
49. Chung, S.W.; Park, J.S.; Kim, S.H.; Shin, S.H.; Oh, J.H. Quality of Life After Arthroscopic Rotator Cuff Repair: Evaluation Using SF-36 and an Analysis of Affecting Clinical Factors. *Am. J. Sport. Med.* **2012**, *40*, 631–639. [[CrossRef](#)] [[PubMed](#)]
50. Frangiamore, S.; Dornan, G.J.; Horan, M.P.; Mannava, S.; Fritz, E.M.; Hussain, Z.B.; Moatshe, G.; Godin, J.A.; Pogorzelski, J.; Millett, P.J. Predictive Modeling to Determine Functional Outcomes after Arthroscopic Rotator Cuff Repair. *Am. J. Sport. Med.* **2020**, *48*, 1559–1567. [[CrossRef](#)] [[PubMed](#)]
51. Razmjou, H.; Holtby, R.; Myhr, T. Gender Differences in Quality of Life and Extent of Rotator Cuff Pathology. *Arthroscopy* **2006**, *22*, 57–62. [[CrossRef](#)]
52. Gasbarro, G.; Ye, J.; Newsome, H.; Jiang, K.; Wright, V.; Vyas, D.; Irrgang, J.J.; Musahl, V. Morphologic Risk Factors in Predicting Symptomatic Structural Failure of Arthroscopic Rotator Cuff Repairs: Tear Size, Location, and Atrophy Matter. *Arthroscopy* **2016**, *32*, 1947–1952. [[CrossRef](#)]
53. Carbone, S.; Gumina, S.; Arceri, V.; Campagna, V.; Fagnani, C.; Postacchini, F. The Impact of Preoperative Smoking Habit on Rotator Cuff Tear: Cigarette Smoking Influences Rotator Cuff Tear Sizes. *J. Shoulder Elb. Surg.* **2012**, *21*, 56–60. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.