

**equinox**<sup>®</sup>  
with Active Intelligence<sup>®</sup>

PREDICT+<sup>™</sup>





## Predict Patient-Specific Outcomes with Exactech's Predict+

Exactech introduces Predict+™, a data-driven, clinical decision support tool that creates personalized patient outcome predictions using preoperative data to anticipate patients' post-operative results after anatomic total shoulder arthroplasty (aTSA) or reverse total shoulder arthroplasty (rTSA).

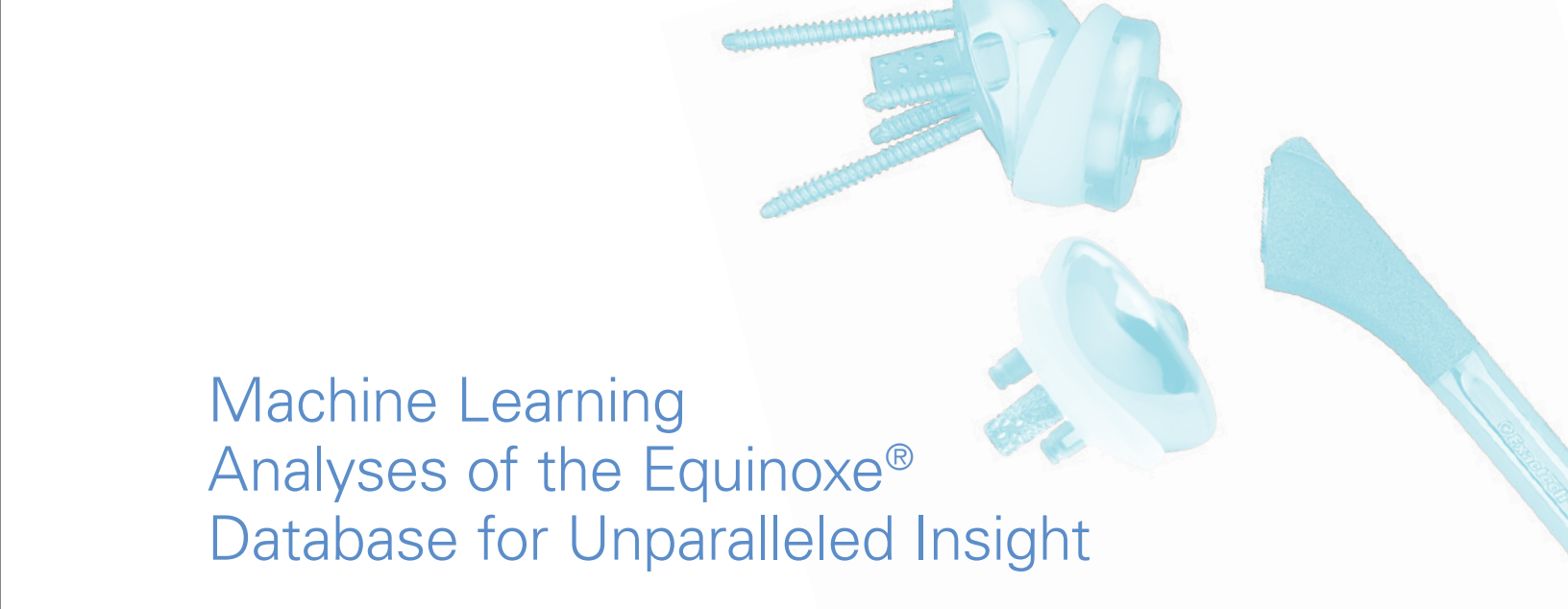
**Predict+ is a new application of clinical research that represents a significant advancement that brings new insights to the patient consultation process. Using machine learning analyses, Predict+ delivers personalized, evidence-based predictions that objectively quantify the potential risk and benefit that an individual patient may experience after either aTSA or rTSA.**

Predict+ is designed to inform and support surgeon clinical treatment decision-making by quantifying precisely how an individual patient will benefit from shoulder arthroplasty. Doing so can better align patient and surgeon

expectations on the magnitude and rate of clinical improvement that they will experience after either aTSA or rTSA. By leveraging the experiences of previous aTSA and rTSA patients with similar demographics, clinical history, diagnoses and comorbidities, Predict+ can preoperatively predict improvements in pain, function and active range of motion that an individual patient may experience at multiple post-operative timepoints. Grounding surgeon and patient expectations in accurate, personalized projections may result in improved patient satisfaction after shoulder arthroplasty.



**PREDICT+** helps power Active Intelligence®, Exactech's unique platform of technologies that helps surgeons engage with patients and peers, solve challenges with predictive tools and optimize the way they perform. Personalized to each surgeon, each patient and each procedure, Active Intelligence helps surgeons make smart decisions throughout the entire journey of care.



# Machine Learning Analyses of the Equinoxe<sup>®</sup> Database for Unparalleled Insight

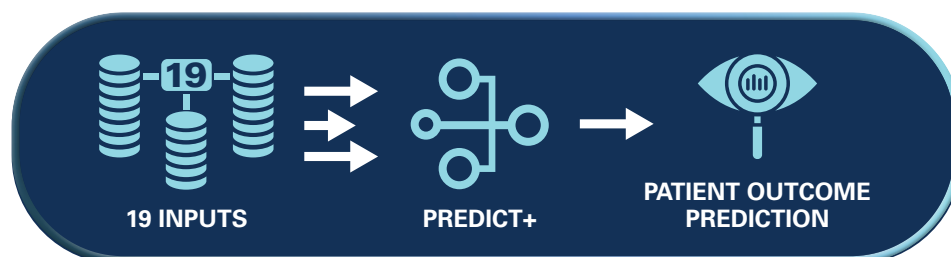
Exactech has built an unrivaled clinical database of a single shoulder prosthesis, consisting of more than 15,000 patients at 40 clinical sites. Exactech is performing machine learning analyses on this clinical database in order to revolutionize patient consultation and care.<sup>1-2,18-21</sup>



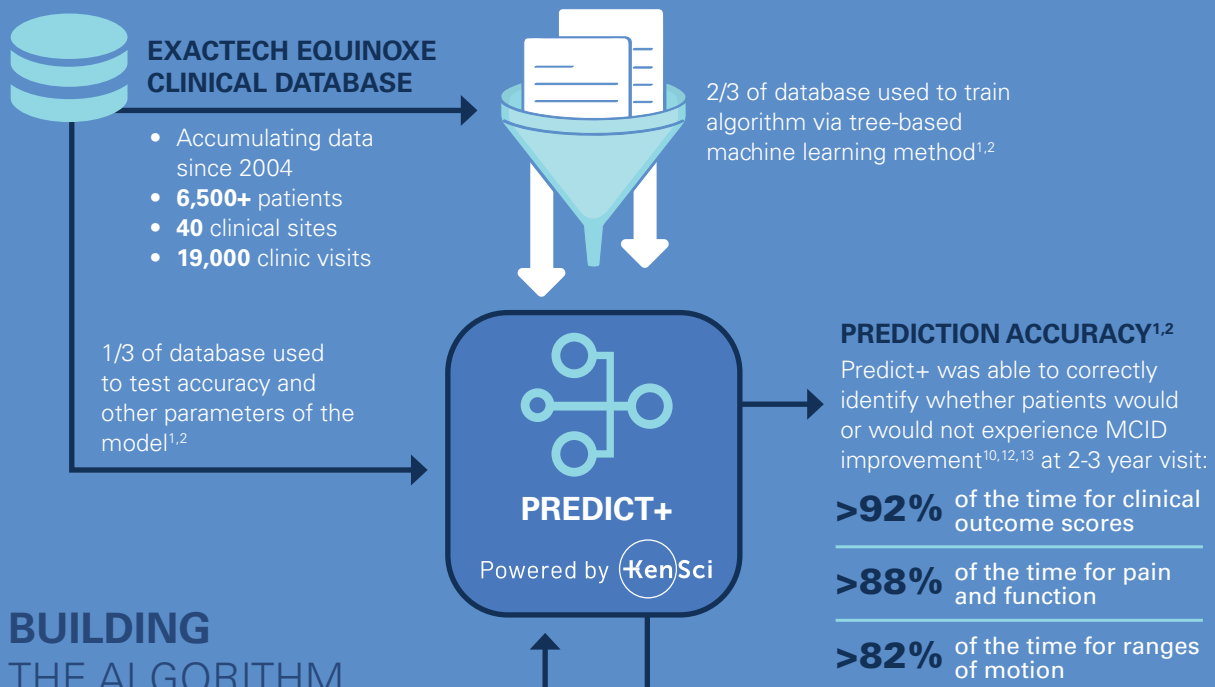
**MACHINE LEARNING** (ML) and Predict+ are examples of Artificial Intelligence, which is intended to enhance and support human decision-making, not replace it. ML is used in other daily applications such as suggesting movies based on viewers' previous choices (e.g. Netflix) or products based on searches and preferences (e.g. Amazon).

Predict+ uses machine learning to identify otherwise hard-to-detect relationships between inputs (like patient demographics and comorbidities) and outputs (like post-operative pain or range of motion). Machine learning is especially effective at finding mathematical relationships that will predict outcomes given a discrete set of inputs. Exactech has partnered with KenSci, a healthcare machine learning company, to analyze our database and create algorithms that can help surgeons predict outcomes for their patients before they even enter the O.R.

The Predict+ algorithm, running on the KenSci AI platform, requires only 19 preoperative datapoints to create predictions for patient-reported outcomes, pain and ROM within seconds.<sup>1,2,6</sup> Anticipated outcomes for your patient are displayed on a graphical dashboard, alongside top factors driving those predicted outcomes. Since prosthesis type also affects outcomes, complications and recovery rate,<sup>6-11</sup> predictions for both aTSA and rTSA are displayed for comparison and also relative to age, gender and prosthesis matched cohorts from our clinical database.



# PREDICT+ MACHINE LEARNING RESEARCH METHODS



## BUILDING THE ALGORITHM

## USING ITS FEATURES

**PATIENT-SPECIFIC INFORMATION**  
**19** preoperative user inputs including:

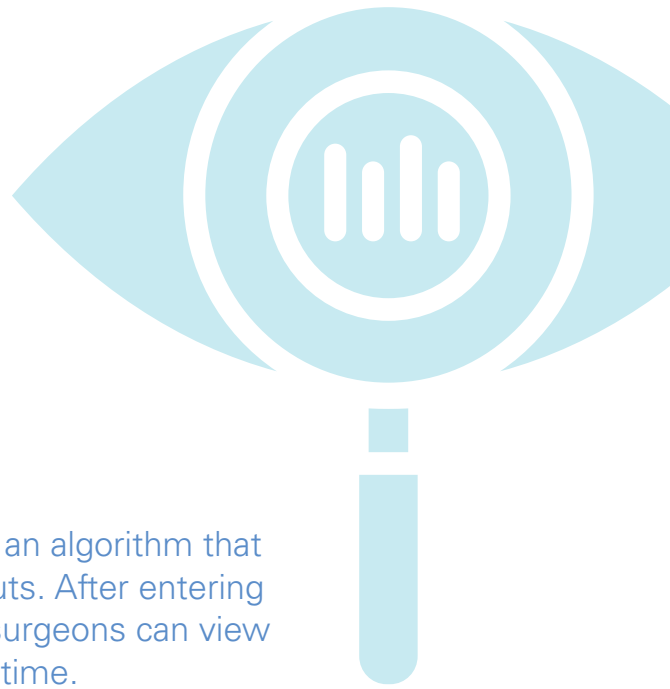
- Demographic data
- Diagnosis and comorbidities
- Patient-reported pain and function
- Ranges of motion

**PATIENT-SPECIFIC RESULTS**  
Dashboard in GPS Web displays helpful information for surgeons to incorporate into their patient counseling and surgical planning:

- Expected scores for **Global Shoulder Function, VAS Pain, active abduction, active forward elevation, internal rotation, Shoulder Arthroplasty Smartscore and active external rotation**, at 7 timepoints from pre-op to 3 to 6 months up to an average of 7 years after surgery.<sup>18-21</sup> With additional inputs, **ASES and Constant** can be predicted for the same 7 timepoints.
- Time post-op at which patient could expect meaningful (MCID<sup>12</sup>) and significant (SCB<sup>13</sup>) recovery
- Comparison of this patient's anticipated outcomes to an age- and gender-matched cohort
- Comparison of this patient's anticipated outcomes for aTSA and rTSA implants
- Complication risks for patients in age- and gender-matched cohort<sup>6</sup>
- Patient-specific factors affecting predictions, which may be modifiable

**Minimal Clinically Important Difference (MCID):** The minimal difference in a clinical outcome measure that a patient perceives as beneficial and meaningful, e.g. "better after surgery."

**Substantial Clinical Benefit (SCB):** A substantial difference in a clinical outcome measure, e.g. "much better after surgery."



# Machine Learning for Predicting Patient-Specific Outcomes

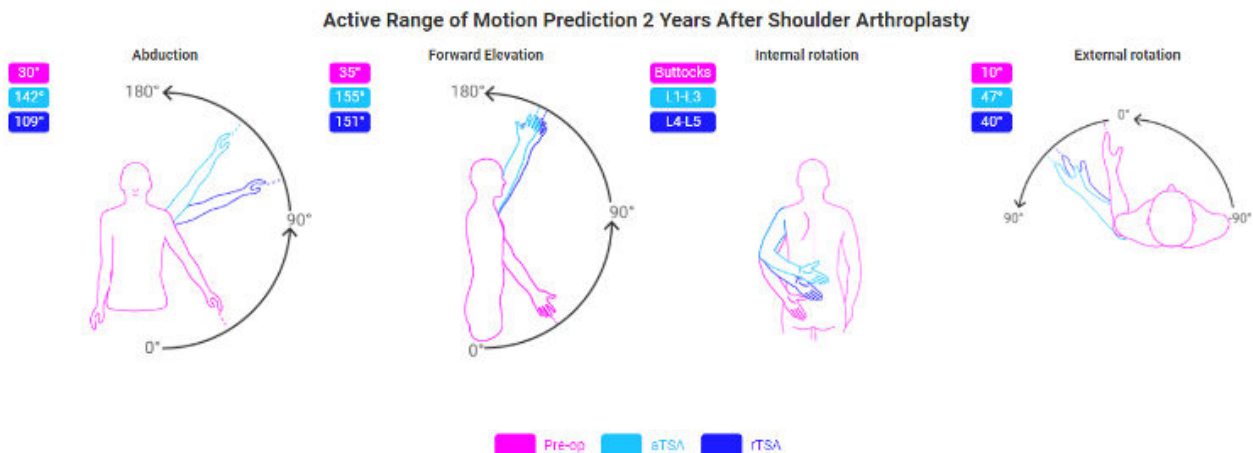
Using machine learning, Exactech has developed an algorithm that predicts patient outcomes from preoperative inputs. After entering patient-specific data into the Predict+ software, surgeons can view expected pain relief and functional recovery over time.

To train the Predict+ algorithm, we analyzed our clinical database consisting of 291 inputs and identified those parameters that most strongly influenced the outputs surgeons and patients care about, like pain and functional recovery.<sup>1</sup> Then, using 19 most influential inputs (i.e. the minimal feature set), we trained the Predict+ model using 2/3 of a database of more than 6,500+ patients and more than 19,000 follow-up visits in order to create a model that

predicts 7 different aTSA and rTSA outcomes at 7 timepoints.<sup>2</sup> Using the remaining 1/3 of the database, the predictions were tested. With our minimal feature set 19-input model, we can predict with 82-99% (AUROC = 0.70-0.95) accuracy when a patient will achieve MCID improvements in outcomes (depending on the specific outcome measure) and with 76-90% accuracy (AUROC = 0.70-0.90) when a patient will achieve SCB improvements in outcomes.<sup>20</sup>

## Area Under the Receiver Operating Curve (AUROC): How good are the predictions?

In machine learning, Area Under the Receiver Operating Curve (AUROC) is a measure of a predictive algorithm's sensitivity and specificity, and values  $>0.7$  are considered acceptable,  $>0.8$  good and  $>0.9$  excellent discrimination for a predictive model. Minimal feature set AUROCs ranged from 0.70-0.95, retaining acceptable to excellent sensitivity and specificity for most measures.

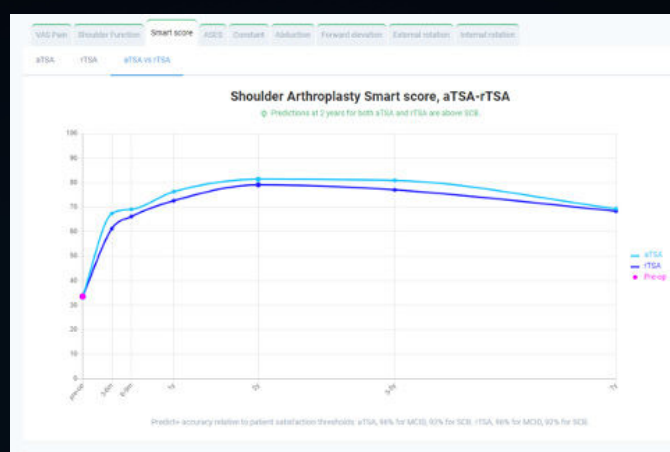
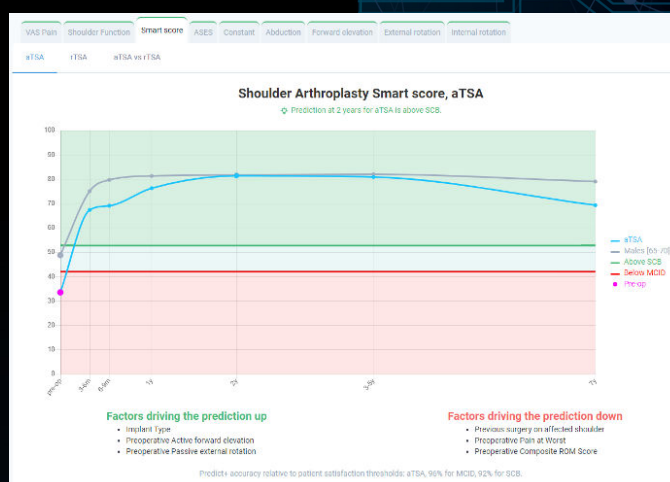


# Defining Successful Outcomes

Understanding and defining patient expectations prior to elective surgery plays a critical role in optimizing outcomes. However, patients experience different clinical outcomes depending on their age, sex and BMI.<sup>10,14-17</sup> Appropriately establishing patient expectations that account for these demographic and other preoperative differences is critical to improve patient satisfaction.

To support surgeon and patient decision-making and improve patient satisfaction, Exactech has defined levels of improvement in outcomes that are clinically relevant.<sup>10,12,13</sup>

Exactech has determined MCID and SCB for patients over time, stratified by implant type, patient age and sex.<sup>10,12,13</sup> These values for MCID and SCB are displayed on each patient's Predict+ dashboard in order to visualize when "better" and "much better" patient satisfaction-based improvement thresholds could be reasonably expected.



## Scientific & Clinical Evidence

1. What is the Accuracy of Three Different Machine Learning Techniques to Predict Clinical Outcomes after Shoulder Arthroplasty?.....8	12. Quantifying Success after Total Shoulder Arthroplasty: The Minimal Clinically Important Difference.....19
2. Using Machine Learning to Predict Clinical Outcomes after Shoulder Arthroplasty with a Minimal Feature Set.....9	13. Quantifying Success after Total Shoulder Arthroplasty: The Substantial Clinical Benefit.....20
3. Preoperative External Rotation Deficit Does Not Predict Poor Outcomes or Lack of Improvement after Reverse Total Shoulder Arthroplasty ..... 10	14. The Effect of BMI on Internal Rotation and Function Following Anatomic and Reverse Total Shoulder Arthroplasty.....21
4. Patient-Reported Outcomes for Reverse Total Shoulder Arthroplasty: A Comparative Risk Factor Analysis of Improved Versus Unimproved Cases 11	15. Results of Total Shoulder Arthroplasty in Patients Aged 55 Years or Younger Versus Those Older Than 55 Years: An Analysis of 1135 Patients with Over 2 Years of Follow-Up.....22
5. Preoperative Parameters that Predict Postoperative Patient Reported Outcome Measures and Range of Motion with Anatomic and Reverse Total Shoulder Arthroplasty ..... 12	16. Early Outcomes of Shoulder Arthroplasty According to Sex.....23
6. Comparison of Complication Types and Rates Associated with Anatomic and Reverse Total Shoulder Arthroplasty..... 13	17. Effects of Body Mass Index on Outcomes in Total Shoulder Arthroplasty.....24
7. Comparison of Survivorship and Performance of a Platform Shoulder System in Anatomic and Reverse Total Shoulder Arthroplasty ..... 14	18. Using Machine Learning to Predict Internal Rotation after Anatomic and Reserve Total Shoulder Arthroplasty.....25
8. Deltoid Fatigue: A Longitudinal Assessment of Reverse Shoulder Arthroplasty over Time ..... 15	19. Validation of a Machine Learning Derived Clinical Metric to Quantify Outcomes after TSA.....26
9. Defining the Tipping Point for Primary Shoulder Arthroplasty..... 16	20. Development of a Predictive Model for a Machine Learning Derived Shoulder Arthroplasty Clinical Outcome Score .....27
10. Are Age and Patient Gender Associated with Different Rates and Magnitudes of Clinical Improvement after Reverse Shoulder Arthroplasty? ..... 17	21. Use of Machine Learning to Assess the Predictive Value of 3 Commonly Used Clinical Measures to Quantify Outcomes After Total Shoulder Arthroplasty.....28
11. Rate of Improvement in Clinical Outcomes with Anatomic and Reverse Total Shoulder Arthroplasty ..18	22. References.....29

# What is the Accuracy of Three Different Machine Learning Techniques to Predict Clinical Outcomes after Shoulder Arthroplasty?

Kumar V, Roche C, Overman S, Simovitch R, Flurin PH, Wright T, Zuckerman J, Routman H, Teredesai A. *Clin Orthop Relat Res.* 2020 Apr 20.

## BACKGROUND

Machine learning techniques can identify complex relationships in large healthcare datasets and build prediction models that better inform physicians in ways that can assist in patient treatment decision-making. In the domain of shoulder arthroplasty, machine learning appears to have the potential to anticipate patients' results after surgery, but this has not been well explored.

### Questions/purposes:

1. What is the accuracy of machine learning to predict the American Shoulder and Elbow Surgery (ASES), University of California Los Angeles (UCLA), Constant, global shoulder function, and VAS pain scores, as well as active abduction, forward flexion, and external rotation at 1 year, 2 to 3 years, 3 to 5 years, and more than 5 years after anatomic total shoulder arthroplasty (aTSA) or reverse total shoulder arthroplasty (rTSA)?
2. What is the accuracy of machine learning to identify whether a patient will achieve clinical improvement that exceeds the minimal clinically important difference (MCID) threshold for each outcome measure?
3. What is the accuracy of machine learning to identify whether a patient will achieve clinical improvement that exceeds the substantial clinical benefit threshold for each outcome measure?

## METHODS

A machine learning analysis was conducted on a database of 7811 patients undergoing shoulder arthroplasty of one prosthesis design to create predictive models for multiple clinical outcome measures. Excluding patients with revisions, fracture indications, and hemiarthroplasty resulted in 6210 eligible primary aTSA and rTSA patients, of whom 4782 patients with 11,198 postoperative follow-up visits had sufficient preoperative, intraoperative, and postoperative data to train and test the predictive models. Preoperative clinical data from 1895 primary aTSA patients and 2887 primary rTSA patients were analyzed using three

commercially available supervised machine learning techniques: linear regression, XGBoost, and Wide and Deep, to train and test predictive models for the ASES, UCLA, Constant, global shoulder function, and VAS pain scores, as well as active abduction, forward flexion, and external rotation. Our primary study goal was to quantify the accuracy of three machine learning techniques to predict each outcome measure at multiple postoperative timepoints after aTSA and rTSA using the mean absolute error between the actual and predicted values. Our secondary study goals were to identify whether a patient would experience clinical improvement greater than the MCID and substantial clinical benefit anchor-based thresholds of patient satisfaction for each outcome measure as quantified by the model classification parameters of precision, recall, accuracy, and area under the receiver operating curve.

## RESULTS

Each machine learning technique demonstrated similar accuracy to predict each outcome measure at each postoperative point for both aTSA and rTSA, though small differences in prediction accuracy were observed between techniques. Across all postsurgical timepoints, the Wide and Deep technique was associated with the smallest mean absolute error and predicted the postoperative ASES score to  $\pm 10.1$  to 11.3 points, the UCLA score to  $\pm 2.5$  to 3.4, the Constant score to  $\pm 7.3$  to 7.9, the global shoulder function score to  $\pm 1.0$  to 1.4, the VAS pain score to  $\pm 1.2$  to 1.4, active abduction to  $\pm 18$  to 21°, forward elevation to  $\pm 15$  to 17°, and external rotation to  $\pm 10$  to 12°. These models also accurately identified the patients who did and did not achieve clinical improvement that exceeded the MCID (93% to 99% accuracy for patient-reported outcome measures (PROMs) and 85% to 94% for pain, function, and ROM measures) and substantial clinical benefit (82% to 93% accuracy for PROMs and 78% to 90% for pain, function, and ROM measures) thresholds.



## CONCLUSIONS

Machine learning techniques can use preoperative data to accurately predict clinical outcomes at multiple postoperative points after shoulder arthroplasty and accurately risk-stratify patients by preoperatively identifying who may and who may not achieve MCID and substantial clinical benefit improvement thresholds for each outcome measure.

## CLINICAL RELEVANCE

Three different commercially available machine learning techniques were used to train and test models that predicted clinical outcomes after aTSA and rTSA; this device-type comparison was performed to demonstrate how predictive modeling techniques can be used in the near future to help answer unsolved clinical questions and augment decision-making to improve outcomes after shoulder arthroplasty.

# Using Machine Learning to Predict Clinical Outcomes after Shoulder Arthroplasty with a Minimal Feature Set

**Kumar V, Roche C, Overman S, Simovitch R, Flurin PH, Wright T, Zuckerman J, Routman H, Teredesai A. J Shoulder Elbow Surg. In Press. 2020.**

## BACKGROUND

A machine learning analysis was conducted on 5,774 shoulder arthroplasty patients to create predictive models for multiple clinical outcome measures after anatomic Total Shoulder Arthroplasty (aTSA) and reverse Total Shoulder Arthroplasty (rTSA). The goal of this study is to compare the accuracy associated with a full feature set predictive model (e.g. full model = 291 parameters) and a minimal feature set model (e.g. abbreviated model = 19 input parameters) to predict clinical outcomes in order to assess the efficacy of using a minimal feature set of inputs as a shoulder arthroplasty clinical decision-support tool.

## METHODS

Clinical data from 2,153 primary aTSA patients and 3,621 primary rTSA patients were analyzed using the XGBoost machine learning technique to create and test predictive models for multiple outcome measures at different post-operative timepoints using a full and abbreviated model. Mean absolute errors (MAE) quantified the difference between actual and predicted outcomes, and each model also predicted if a patient would experience clinical improvement greater than the minimal clinically important difference (MCID) and substantial clinical benefit (SCB) patient satisfaction anchor-based thresholds for each outcome measure at 2-3 years after surgery.

## RESULTS

Across all post-operative timepoints analyzed, the full and abbreviated models had similar MAE for the American Shoulder and Elbow Surgeons (ASES) (full model =  $\pm 11.7$  vs. abbreviated model =  $\pm 12.0$ ), Constant ( $\pm 8.9$  vs.  $\pm 9.8$ ), Global Shoulder Function ( $\pm 1.4$  vs.  $\pm 1.5$ ), Visual Analog Scale (VAS) pain ( $\pm 1.3$  vs.  $\pm 1.4$ ), active abduction ( $\pm 20.4^\circ$  vs.  $\pm 21.8^\circ$ ), forward elevation ( $\pm 17.6^\circ$  vs.  $\pm 19.2^\circ$ ), and external rotation ( $\pm 12.2^\circ$  vs.  $\pm 12.6^\circ$ ). Marginal improvements in MAE were observed for each outcome measure prediction when the abbreviated model was supplemented with implant size/type data and measurements of native glenoid anatomy. The full and abbreviated models each effectively risk-stratified patients using only pre-operative data by accurately identifying patients with improvement greater than the MCID and SCB thresholds.

## DISCUSSION

Our study demonstrated the full and abbreviated machine learning models achieved similar accuracy to predict clinical outcomes after aTSA and rTSA at multiple post-operative timepoints. These promising results demonstrate an efficient utilization of machine learning algorithms to predict clinical outcomes. The use of a minimal feature set of only 19 preoperative inputs suggests that this tool may be easily used during a surgical consultation to improve decision-making related to shoulder arthroplasty.



# Preoperative External Rotation Deficit Does Not Predict Poor Outcomes or Lack of Improvement after Reverse Total Shoulder Arthroplasty

---

**M Parsons, HD Routman, CP Roche, RJ Friedman**  
**Journal of Orthopaedics. Volume 21, September–October 2020, Pages 379-383.**

## INTRODUCTION

The purpose was to compare postoperative outcomes and functional improvement between patients with preoperative aER deficits vs. preserved aER function.

## RESULTS

There were 115 patients in the  $<0^\circ$  aER group and 314 in the  $\geq 30^\circ$  aER group. Preoperative patients in the  $<0^\circ$  group were worse for all measures except subjective pain while post-operatively, they had significantly greater improvement for all measures of motion. Postoperatively, both groups achieved comparable scores for forward elevation, pain, SST and ASES.

## CONCLUSION

This study demonstrates that patients with a complete aER deficit can recover substantial and comparable function after RTSA.

# Patient-Reported Outcomes for Reverse Total Shoulder Arthroplasty: A Comparative Risk Factor Analysis of Improved Versus Unimproved Cases

**Moby Parsons, Howard D Routman, Christopher P Roche, Richard J Friedman**  
**JSES Open Access. 2019 Sep 13;3(3):174-178.**

## BACKGROUND

The purpose of this study was to compare characteristics of patients who reported to be subjectively unimproved vs. improved after reverse total shoulder arthroplasty.

## METHODS

Data were derived from a prospective registry of patients who underwent reverse total shoulder arthroplasty with a minimum 2-year follow-up. Patients were asked to rate their subjective satisfaction and then divided into those who were unchanged or worse (unimproved group [UG]) vs. better or much better (improved group [IG]). The groups were compared for differences in demographic characteristics, preoperative factors, functional outcomes, and complications.

## RESULTS

There were 1425 patients in the IG and 134 patients in the UG. Patients in the IG were more likely to have a diagnosis of osteoarthritis. Patients in the UG were more likely to have coronary artery disease and diabetes and to have undergone prior surgery. No differences in implant configuration were found between groups. Preoperative measures for patients in the UG were worse for pain and function but not for range of motion. The outcomes in patients in the UG were worse for all postoperative measures, as well as for preoperative-to-postoperative improvement. Of the patients in the UG, 48% continued to have moderate to severe pain postoperatively. The complication rate was significantly higher in the UG.

## DISCUSSION

Up to 8.5% of patients rate themselves as unimproved after surgery. These patients are more likely to have certain comorbidities and to have undergone prior surgery. Although outcomes were significantly worse for all measures in the UG, improvement occurred in all measures despite patients subjectively being worse or unchanged. Residual pain and difficulty sleeping play a substantial role in subjective assessment of overall outcome.

# Preoperative Parameters that Predict Postoperative Patient Reported Outcome Measures and Range of Motion with Anatomic and Reverse Total Shoulder Arthroplasty

**Richard J Friedman, Josef Eichinger, Bradley Schoch, Thomas Wright, Joseph Zuckerman, Pierre-Henri Flurin, Charlotte Bolch, Chris Roche. JSES Open Access. In Press. 2019.**

## BACKGROUND

Preoperative factors that most influence postoperative outcomes of both anatomic total shoulder arthroplasty (aTSA) and reverse total shoulder arthroplasty (rTSA) are unknown. The purpose of this study was to identify the preoperative parameters that significantly influence postoperative outcomes of aTSA and rTSA.

## METHODS

The outcomes of 1089 aTSA patients and 1332 rTSA patients (mean follow-up period, 49 months) from an international registry with a single platform system were analyzed. A multiple linear regression model with backward stepwise selection identified the preoperative parameters that were significant predictors of postoperative clinical outcome metric scores and motion measures for both rTSA and aTSA.

## RESULTS

For both aTSA and rTSA patients, numerous preoperative parameters that influence postoperative outcomes were identified. Greater postoperative range of motion (ROM) was significantly influenced by greater preoperative ROM. For aTSA, greater postoperative American Shoulder and Elbow Surgeons (ASES) scores were significantly influenced by greater preoperative ASES scores, no history of shoulder surgery, and the presence of greater preoperative active external rotation. For rTSA, greater postoperative ASES scores were significantly influenced by greater preoperative ASES scores, no history of shoulder surgery, no history of tobacco use, less preoperative passive external rotation, and greater preoperative active external rotation.

## CONCLUSIONS

This study quantified the preoperative predictors of postoperative clinical outcome metric scores and ROM for both aTSA and rTSA. Numerous significant associations were identified, including demographic and comorbidity risk factors. These associations may be helpful for surgeons to consider when counseling patients regarding aTSA versus rTSA and to establish more accurate expectations prior to surgery.

# Comparison of Complication Types and Rates Associated with Anatomic and Reverse Total Shoulder Arthroplasty

**Stephen A Parada, Pierre-Henri Flurin, Thomas W. Wright, Joseph D.Zuckerman, Josie A.Elwell, Christopher P.Roche, Richard J.Friedman J Shoulder Elbow Surg. In Press. 2020.**

## INTRODUCTION

Complications after anatomic (aTSA) and reverse (rTSA) total shoulder arthroplasty can be devastating to a patient's quality of life and require revisions which are costly to both the patient and the health care system. The purpose of this study is to determine the types, incidence and timing of complications following aTSA and rTSA using an international database of patients who received a single platform total shoulder arthroplasty system, in order to quantify the types of failure modes and the differences that occur between aTSA and rTSA.

## METHODS


2224 aTSA (1090M/1134F) and 4158 rTSA (1478M/2680F) patients were enrolled in an international database of primary shoulder arthroplasty performed by 40 different surgeons in the US/Europe. Adverse events and revisions reported for these 6382 patients were analyzed to identify the most common failure modes associated for both aTSA and rTSA.

## RESULTS

Of 2224 aTSA patients, 239 adverse events were reported for a complication rate of 10.7% and 124 revisions for a revision rate of 5.6%. The top three complications for aTSA were rotator cuff tear/subscapularis failure (n=69, complication rate=3.1%, revision rate=1.9%), aseptic glenoid loosening (n=55, complication rate=2.5%, revision rate=1.9%) and infection (n=28, complication rate=1.3%, revision rate=0.8%). Of 4158 rTSA patients, 372 adverse events were reported for a complication rate of 8.9% and 104 revisions for a revision rate of 2.5%. The top three complications for rTSA were acromial/scapular fracture/pain (n=102, complication rate=2.5%, revision rate=0.0%), instability (n=60, complication rate=1.4%, revision rate=1.0%) and pain (n=49, complication rate=1.2%, revision rate=0.2%).

## CONCLUSIONS

This large database analysis quantified complication and revision rates for aTSA and rTSA. We found aTSA and rTSA complication rates of 10.7% and 8.9%, respectively; with revision surgery rates of 5.6% and 2.5%, respectively. The two most common complications for each prosthesis type (aTSA: subscapularis/rotator cuff tears, aseptic glenoid loosening; rTSA: acromial/scapular fractures, instability) were unique to each device. The rate of infection was similar for both. Future prosthesis and technique development should work to mitigate these common complication types in order to reduce their rate of occurrence.



# Comparison of Survivorship and Performance of a Platform Shoulder System in Anatomic and Reverse Total Shoulder Arthroplasty

---

**Pierre Henri Flurin, Carl Tams, Ryan W. Simovitch, Christopher Knudsen, Christopher Roche, Thomas Wright, Joseph Zuckerman, Bradley S. Schoch JSES International. 2020. In press.**

## BACKGROUND

Contemporary studies note sustained clinical benefit and decreasing complications following reverse total shoulder arthroplasty (RTSA), which warrant a comparison to the standard anatomic total shoulder arthroplasty (ATSA). The purpose of this study is to evaluate and compare differences in mid-term survivorship between ATSA and RTSA patients treated with a single platform shoulder prosthesis. Secondary objectives include a comparison of the clinical outcomes and complication profile for each procedure.

## METHODS

A prospective analysis of all primary ATSA and RTSA performed by three surgeons between 2007-2012 was conducted. Selection of the ATSA or RTSA implant configuration was determined by the surgeon per their clinical understanding of each individual patient's glenoid morphology, rotator cuff, and patient expectations. All 778 procedures were performed using a single platform shoulder system.

## RESULTS

Survivorship for both ATSA and RTSA were similar at all time points; ATSA at 2 and 8 years was 98.5% and 96.0% while RTSA at 2 and 8 years was 98.7% and 96.0% [p=0.392]. All postoperative range of motions scores were greater for ATSA patients than for RTSA patients. The overall rate of complications between the ATSA and RTSA groups was similar [6.3% vs 4.9%, p=0.414].

## CONCLUSIONS

Based on this cohort comparison, both ATSA and RTSA demonstrated similar survivorship at 8 years after surgery with multiple surgeons practicing in different countries. Our results demonstrate that the RTSA and ATSA implants have comparable results and can be expected to provide similar implant longevity over the mid-term with excellent functional outcomes.

# Deltoid Fatigue: A Longitudinal Assessment of Reverse Shoulder Arthroplasty over Time

**Bradley S Schoch, Marie Vigan, Christopher Roche, Moby Parsons, Thomas W Wright, Joseph J King, Jean David Werthel. J Shoulder Elbow Surg. In press. 2020.**

## BACKGROUND

Studies evaluating the mid-term performance of reverse shoulder arthroplasty (RSA) have identified a drop in the Constant-Murley score between 6 and 8 years after surgery, which is most affected by a loss of forward elevation and strength. Alterations of the deltoid length and moment arm after RSA leads to non-physiologic stress on the deltoid muscle. Concern has arisen that the long-term implications of increased deltoid work may be causing “deltoid fatigue.” The purpose of this study is to evaluate the long-term effects of RSA on overhead range of motion (ROM) and to validate the hypothesis of “deltoid fatigue.”

## METHODS

A retrospective review of 165 RSA over a 5 year period was performed. Diagnoses were limited to cuff tear arthropathy (CTA), osteoarthritis with rotator cuff deficiency (OA), and irreparable rotator cuff tear (RCT). All procedures were performed using a single implant system. Patients were evaluated longitudinally at multiple time points. Patients were required to have a minimum of three follow-up visits, with at least one exceeding 5 years. ROM and patient reported outcome measures (PROM) were evaluated using linear-mixed models for repeated measures to evaluate changes in outcome measures over time. A secondary analysis was performed to assess the influence of patient demographic factors on observed changes in ROM and PROMs.

## RESULTS

Primary RSA shoulders were observed to lose 0.8° of forward elevation and abduction per year starting at 1 year post-operatively ( $p=0.006$ ), without a significant drop at mid-term follow-up. No significant change in external rotation or internal rotation was observed. Males and patients with a diagnosis of OA showed greater baseline overhead ROM at one year post-operatively, but the subsequent rate of functional decline occurred at similar rates regardless of age, gender or indication.

## DISCUSSION

This study challenges the previous theory of “deltoid fatigue” resulting in a significant loss of overhead ROM beginning 6 to 8 years after index arthroplasty. However, a slower progressive decline in overhead ROM in well-functioning RSA shoulders was observed, averaging 0.8 degrees of overhead ROM per year. This progressive deterioration occurs at a slightly greater rate than the that observed in the natural shoulder. The observed rate of functional decline was found to be independent of age, gender and preoperative diagnosis.



# Defining the Tipping Point for Primary Shoulder Arthroplasty

---

**Bradley S.Schoch, Joseph J.King, Thomas W.Wright, MarieVigan, Jean David Werthel**  
**JSES Open Access. 2019. Volume 3, Issue 4, December 2019, Pages 273-277.**

## BACKGROUND

Although risk factors for poor outcomes and complications have been studied, there remain limited objective criteria to guide surgeons about the timing of arthroplasty. The purpose of this study was to further characterize the tipping-point scores for a group of patient-reported outcome measures (PROMs) in patients undergoing primary shoulder arthroplasty.

## METHODS

We retrospectively reviewed 5670 primary shoulder arthroplasties (1833 anatomic total shoulder arthroplasties and 3837 reverse shoulder arthroplasties [RSAs]) performed over a 10-year period. Preoperative range of motion, PROMs (American Shoulder and Elbow Surgeons, Simple Shoulder Test, and Shoulder Pain and Disability Index scores), and Constant scores were evaluated. The tipping point for each PROM was evaluated. Univariate and multivariate analyses were performed to assess risk factors for lower tipping points.

## RESULTS

Patients undergoing RSA demonstrated lower tipping points for all range-of-motion parameters as well as American Shoulder and Elbow Surgeons, Shoulder Pain and Disability Index, and Simple Shoulder Test scores. Female sex was predictive of a lower tipping point prior to shoulder arthroplasty, regardless of implant type. When the total shoulder arthroplasty subgroup was evaluated, both female sex and a higher body mass index were shown to be associated with a lower tipping point.

## DISCUSSION

The choice to undergo shoulder arthroplasty is a multifactorial decision that encompasses both physical and social factors. Female patients and patients undergoing RSA are more likely to accept slightly worse shoulder function prior to making the decision to undergo shoulder arthroplasty.



# Are Age and Patient Gender Associated with Different Rates and Magnitudes of Clinical Improvement after Reverse Shoulder Arthroplasty?

**Richard J Friedman, Emilie V Cheung, Pierre-Henri Flurin, Thomas Wright, Ryan W Simovitch, Charlotte Bolch, Christopher P Roche, Joseph D Zuckerman** *Clin Orthop Relat Res.* 2018 Jun;476(6):1264-1273.

## BACKGROUND

An improved understanding of how gender differences and the natural aging process are associated with differences in clinical improvement in outcome metric scores and ROM measurements after reverse total shoulder arthroplasty (rTSA) may help physicians establish more accurate patient expectations for reducing postoperative pain and improving function.

## QUESTIONS/PURPOSES

(1) Is gender associated with differences in rTSA outcome scores like the Simple Shoulder Test (SST), the UCLA Shoulder score, the American Shoulder and Elbow Surgeons (ASES) Shoulder score, the Constant Shoulder score, and the Shoulder Pain and Disability Index (SPADI) and ROM? (2) Is age associated with differences in rTSA outcome scores and ROM? (3) What factors are associated with the combined interaction effect between age and gender? (4) At what time point during recovery does most clinical improvement occur, and when is full improvement reached?

## METHODS

We quantified and analyzed the outcomes of 660 patients (424 women and 236 men; average age,  $72 \pm 8$  years; range, 43-95 years) with cuff tear arthropathy or osteoarthritis and rotator cuff tear who were treated with rTSA by 13 shoulder surgeons from a longitudinally maintained international database using a linear mixed effects statistical model to evaluate the relationship between clinical improvements and gender and patient age. We used five outcome scoring metrics and four ROM assessments to evaluate clinical outcome differences.

## RESULTS

When controlling for age, men had better SST scores (mean difference [MD] = 1.41 points [95% confidence interval {CI}, 1.07-1.75],  $p < 0.001$ ), UCLA scores (MD = 1.76 [95% CI, 1.05-2.47],  $p < 0.001$ ), Constant scores (MD = 6.70 [95% CI, 4.80-8.59],  $p < 0.001$ ), ASES scores (MD = 7.58 [95% CI, 5.27-9.89],  $p < 0.001$ ), SPADI scores (MD = -12.78 [95% CI, -16.28 to -9.28],  $p < 0.001$ ), abduction (MD =  $5.79^\circ$  [95% CI, 2.74-8.84],  $p < 0.001$ ), forward flexion (MD =  $7.68^\circ$  [95% CI, 4.15-11.20],  $p < 0.001$ ), and passive external rotation (MD =  $2.81^\circ$  [95% CI, 0.81-4.8],  $p = 0.006$ ). When controlling for gender, each 1-year increase in age was associated with an improved ASES score by 0.19 points (95% CI, 0.04-0.34,  $p = 0.011$ ) and an improved SPADI score by -0.29 points (95% CI, -0.46 to 0.07,  $p = 0.020$ ). However, each 1-year increase in age was associated with a mean decrease in active abduction by  $0.26^\circ$  (95% CI, -0.46 to 0.07,  $p = 0.007$ ) and a mean decrease of forward flexion by  $0.39^\circ$  (95% CI, -0.61 to 0.16,  $p = 0.001$ ). A combined interaction effect between age and gender was found only with active external rotation: in men, younger age was associated with less active external rotation and older age was associated with more active external rotation ( $\beta_0$  [intercept] = 11.029,  $\beta_1$  [slope for age variable] = 0.281,  $p = 0.009$ ). Conversely, women achieved no difference in active external rotation after rTSA, regardless of age at the time of surgery ( $\beta_0$  [intercept] = 34.135,  $\beta_1$  [slope for age variable] = -0.069,  $p = 0.009$ ). Finally, 80% of patients achieved full clinical improvement as defined by a plateau in their outcome metric score and 70% of patients achieved full clinical improvement as defined by a plateau in their ROM measurements by 12 months followup regardless of gender or patient age at the time of surgery with most improvement occurring in the first 6 months after rTSA.

## CONCLUSIONS

Gender and patient age at the time of surgery were associated with some differences in rTSA outcomes. Men had better outcome scores than did women, and older patients had better outcome scores but smaller improvements in function than did younger patients. These results demonstrate rTSA outcomes differ for men and women and for different patient ages at the time of surgery, knowledge of these differences, and also the timing of improvement plateaus in outcome metric scores and ROM measurements can both improve the effectiveness of patient counseling and better establish accurate patient expectations after rTSA.

# Rate of Improvement in Clinical Outcomes with Anatomic and Reverse Total Shoulder Arthroplasty

Simovitch RW, Friedman RJ, Cheung EV, Flurin PH, Wright T, Zuckerman JD, Roche C  
*J Bone Joint Surg Am.* 2017 Nov 1;99(21):1801-1811.

## BACKGROUND

This study quantifies the rate of improvement after anatomic and reverse total shoulder arthroplasty; a better understanding of the rate of improvement associated with each prosthesis type may better establish patient expectations for recovery.

## METHODS

Prospectively collected data on 1,183 patients who underwent either anatomic total shoulder arthroplasty (n = 505) or reverse total shoulder arthroplasty (n = 678) were collected. The Simple Shoulder Test (SST), University of California at Los Angeles (UCLA) Shoulder, American Shoulder and Elbow Surgeons (ASES), Constant, and Shoulder Pain and Disability Index (SPADI) scores, along with range of motion, were recorded preoperatively and at routine postoperative time points. All included patients had a minimum follow-up of 2 years. The rate of improvement of these outcome measures was quantified for patients who underwent anatomic total shoulder arthroplasty and those who underwent reverse total shoulder arthroplasty to compare recovery over time.

## RESULTS

In this study, 3,587 visits by 1,183 patients were analyzed and several differences between prosthesis types were noted. Patients who underwent reverse total shoulder arthroplasty experienced larger improvements in the Constant score and active forward flexion, and patients who underwent anatomic total shoulder arthroplasty

demonstrated better improvement in external rotation compared with patients who underwent reverse total shoulder arthroplasty at nearly all time points. By 72 months, improvement in flexion and abduction decreased for each prosthesis type, but in particular for reverse total shoulder arthroplasty. Full improvement was achieved by 24 months, although the majority of improvement was achieved in the first 6 months, with all 5 scoring metrics following a similar rate of improvement. The ASES, SPADI, and UCLA Shoulder scores closely mirrored each other in the magnitude of improvement, and the SST score demonstrated the largest improvement and the Constant score demonstrated the smallest improvement for both anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty.

## CONCLUSIONS

Both anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty reliably result in improved patient outcomes. However, anatomic total shoulder arthroplasty more reliably improves range of motion, particularly external rotation. Most improvement occurs by 6 months, with some additional improvement up to 2 years for both anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty. Although the indications for anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty are substantially different, in addition to the biomechanical differences, the improvement in outcome scores over time can be expected to be very similar. This study is helpful to patients and health-care providers to establish expectations regarding the rate of recovery after total shoulder arthroplasty.

# Quantifying Success after Total Shoulder Arthroplasty: The Minimal Clinically Important Difference

Simovitch R, Flurin PH, Wright T, Zuckerman JD, Roche CP  
J Shoulder Elbow Surg. 2018 Feb;27(2):298-305.

## BACKGROUND

Knowledge of the minimal clinically important difference (MCID) for different shoulder outcome metrics and range of motion after total shoulder arthroplasty (TSA) can be useful to establish a minimum threshold of improvement that defines successful treatment. This study quantifies how MCID varies with different prosthesis types, patient age, gender, and length of follow-up after TSA.

## METHODS

A total of 466 anatomic TSA (aTSA) and reverse TSA (rTSA) with 2-year minimum follow-up were performed by 13 shoulder surgeons. The MCID for the American Shoulder and Elbow Surgeons, Constant, University of California Los Angeles Shoulder Rating Scale, Simple Shoulder Test, Shoulder Pain and Disability Index, global shoulder function, and visual analog scale for pain scores, as well as active abduction, forward flexion, and external rotation, were calculated for different prosthesis types and patient cohorts using an anchor-based method.

## RESULTS

The anchor-based MCID results were American Shoulder and Elbow Surgeons =  $13.6 \pm 2.3$ , Constant score =  $5.7 \pm 1.9$ , University of California Los Angeles Shoulder Rating Scale =  $8.7 \pm 0.6$ , Simple Shoulder Test score =  $1.5 \pm 0.3$ , Shoulder Pain and Disability Index score =  $20.6 \pm 2.6$ , global shoulder function =  $1.4 \pm 0.3$ , pain visual analog scale =  $1.6 \pm 0.3$ , active abduction =  $7^\circ \pm 4^\circ$ , active forward flexion =  $12^\circ \pm 4^\circ$ , and active external rotation =  $3^\circ \pm 2^\circ$ . Female gender and rTSA were associated with lower MCID values compared with male gender and aTSA patients.

## CONCLUSION

The minimum improvement necessary for patients to achieve a result they believe is clinically meaningful after aTSA and rTSA is nominal and was achieved by at least 80% of the patients. Future endeavors should investigate the influence of different anchor questions on the MCID calculation.

# Quantifying Success after Total Shoulder Arthroplasty: The Substantial Clinical Benefit

**Simovitch R, Flurin PH, Wright T, Zuckerman JD, Roche CP.**  
**J Shoulder Elbow Surg. 2018 May;27(5):903-911.**

## BACKGROUND

An understanding of the substantial clinical benefit (SCB) after total shoulder arthroplasty (TSA) may help to gauge a minimum threshold beyond which a patient perceives his or her outcome as being substantially better. This study quantifies SCB for 7 outcome metrics and active motion measurements after shoulder arthroplasty and determines how these values vary based on prosthesis type, patient age at surgery, sex, and length of follow-up.

## METHODS

A total of 1,568 shoulder arthroplasties with 2-year minimum follow-up were performed by 13 shoulder surgeons and enrolled in a multicenter registry. The SCB for the American Shoulder and Elbow Surgeons Shoulder Assessment, Constant Score, University of California Los Angeles Shoulder Rating Scale, Simple Shoulder Test, Shoulder Pain and Disability Index, global shoulder function, and visual analog scale pain scores, as well as active abduction, flexion, and external rotation were calculated for different patient cohorts using an anchor-based method.

## RESULTS

The anchor-based SCB results were American Shoulder and Elbow Surgeons score,  $31.5 \pm 2.0$ ; Constant Score,  $19.1 \pm 1.7$ ; University of California Los Angeles Shoulder Rating Scale score,  $12.6 \pm 0.5$ ; Simple Shoulder Test score,  $3.4 \pm 0.3$ ; Shoulder Pain and Disability Index score,  $45.4 \pm 2.2$ ; global shoulder function,  $3.1 \pm 0.2$ ; visual analog scale,  $3.2 \pm 0.3$ ; active abduction,  $28.5^\circ \pm 3.1^\circ$ ; active forward flexion,  $35.4^\circ \pm 3.5^\circ$ ; and active external rotation,  $11.7^\circ \pm 1.9^\circ$ . Anatomic TSA patients, male patients, and patients of longer follow-up duration were associated with higher SCB values than female patients, reverse TSA patients, and patients of shorter follow-up duration.

## CONCLUSION

Our analysis demonstrated two-thirds of patients achieved the SCB threshold after TSA. Generally, a change of 30% of the total possible score for each outcome metric approximates or exceeds this SCB threshold.

# The Effect of BMI on Internal Rotation and Function Following Anatomic and Reverse Total Shoulder Arthroplasty

**JK Eichinger, MV Rao, JJ Lin, J Goodloe, V Kothandaraman, WR Barfield, SA Parada, C Roche, RJ Friedman**  
**J Shoulder Elbow Surg. In press. 2020.**

## BACKGROUND

The exact relationship between BMI and IR before and after total shoulder arthroplasty has not been studied to date. The purpose of this study is to determine the effects of BMI on the preoperative and postoperative shoulder range of motion and function in anatomic (aTSA) and reverse total shoulder arthroplasty (rTSA), and specifically how internal rotation (IR) affects patient ability to perform IR-related activities of daily living (ADLs).

## METHODS

Patients from a prospective multicenter international shoulder arthroplasty registry who underwent primary rTSA (n=1171) and primary aTSA (n=883) were scored preoperatively and at latest follow-up (2-10 years, mean = 3 years) using the SST, UCLA, ASES, Constant, and SPADI patient reported outcome measures (PROM). Measured active abduction, forward flexion, internal rotation, and active and passive external rotation were recorded and BMI was evaluated as a predictor of motion and patient-reported outcomes. Patient responses to questions regarding the difficulty level of IR-related ADLs were studied. The relationships between BMI, IR, and ability to perform IR-related ADLs were quantified through ANOVA with post-hoc comparisons by Tukey HSD tests where significance was denoted as  $p < 0.05$ .

## RESULTS

BMI was found to be inversely correlated with IR in patients undergoing both aTSA and rTSA, both preoperatively ( $p < 0.001$  and  $p = 0.002$ ) and postoperatively ( $p < 0.001$  and  $p < 0.001$ ). BMI affected range of motion parameters of forward flexion abduction and external rotation but a lesser extent than that of IR. Non-obese patients demonstrated significantly greater IR than overweight, obese, and morbidly obese patients postoperatively for aTSA ( $p < 0.001$ ). For rTSA, non-obese patients had a significantly greater postoperative IR than obese and morbidly obese patients ( $p < 0.001$  and  $p = 0.011$ , respectively). For both aTSA and rTSA patients, mean IR scores significantly differed between patients reporting normal function versus patients reporting slight difficulty, considerable difficulty, or inability to perform IR related ADLs. Increasing IR demonstrated a significant, positive correlation with all PROM for both aTSA and rTSA patients (Pearson's correlation,  $p < 2.2e-16$ ).

## CONCLUSIONS

BMI is an independent predictor of IR, even when controlling for age, gender, glenosphere size, and subscapularis repair. BMI was inversely correlated with the degree of internal rotation and decreased internal rotation significantly negatively impacted ability to perform IR related ADLs.



# Results of Total Shoulder Arthroplasty in Patients Aged 55 Years or Younger Versus Those Older Than 55 Years: An Analysis of 1135 Patients with Over 2 Years of Follow-Up

---

**Patel RB, Muh S, Okoroha KR, Wright TW, Flurin PH, Roche C, Zuckerman JD.**  
**J Shoulder Elbow Surg. 2019 May;28(5):861-868.**

## BACKGROUND

The results of anatomic total shoulder arthroplasty (TSA) in younger patients have not been clearly elucidated. The purpose of this study was to compare early outcomes after TSA in patients aged 55 years or younger versus patients older than 55 years.

## METHODS

A total of 1135 patients who were treated with TSA for glenohumeral arthritis and had a mean follow-up period of over 4 years were retrospectively reviewed. Etiologies included osteoarthritis (n = 1044), osteonecrosis (n = 35), inflammatory arthritis (n = 34), and post-traumatic arthritis (n = 22). Validated outcome measures, range of motion, and patient satisfaction were recorded. Preoperative and postoperative metrics were compared, and a multivariate analysis was performed to isolate age from sex, body mass index, previous surgery, and diagnosis as independent factors.

## RESULTS

Female patients, patients with a history of surgery, and patients with a diagnosis of osteonecrosis were more likely to undergo TSA when aged 55 years or younger. Both age groups showed similar preoperative range of motion and showed no differences in recorded outcome scores. Postoperatively, patients older than 55 years had slightly greater active abduction ( $P = .004$ ) and internal rotation ( $P = .030$ ). A higher percentage of patients older than 55 years rated their outcome as better or much better compared with those aged 55 years or younger ( $P = .003$ ).

## CONCLUSIONS

Female sex, a history of surgery, and a diagnosis of osteonecrosis were associated with undergoing TSA when aged 55 years or younger. Despite similar preoperative function and minor differences in postoperative range of motion and outcome scores, patients aged 55 years or younger reported lower overall satisfaction with their TSA.

# Early Outcomes of Shoulder Arthroplasty According to Sex

Okoroha KR, Muh S, Gabbard M, Evans T, Roche C, Flurin PH, Wright TW, Zuckerman JD.  
JSES Open Access. 2019 Mar 16;3(1):43-47.

## HYPOTHESIS

Among patients undergoing shoulder arthroplasty (SA), female patients would have worse outcomes than their male counterparts.

## METHODS

A multicenter prospective cohort of 2364 patients (1365 female and 999 male patients) treated with total SA or reverse total SA from 2007 to 2015 was retrospectively analyzed. Results were assessed using several validated outcome measures and range-of-motion testing. A multivariable analysis identified differences in preoperative values, postoperative values, and preoperative-to-postoperative improvements while adjusting for possible confounders.

## RESULTS

The mean follow-up period was  $45.9 \pm 23.7$  months in female patients and  $46.4 \pm 23.6$  months in male patients. Women underwent SA at a significantly older age ( $70.8 \pm 8.4$  years) than men ( $67.6 \pm 8.8$  years,  $P < .01$ ) and began with lower preoperative outcome scores and range-of-motion measurements: American Shoulder and Elbow Surgeons score ( $P < .01$ ), Constant score ( $P < .01$ ), Simple Shoulder Test score ( $P < .01$ ), active abduction ( $P$

$< .01$ ), forward flexion ( $P < .01$ ), and external rotation ( $P = .02$ ). Postoperatively, both groups showed significant improvement. When we evaluated overall improvement from preoperative values, female patients showed increased improvements in the American Shoulder and Elbow Surgeons score ( $P = .04$ ) and Simple Shoulder Test score ( $P < .01$ ), as well as active forward elevation ( $P < .01$ ) and external rotation ( $P = .02$ ). However, the difference in improvements did not reach the minimal clinically important difference. Women had a higher incidence of component loosening ( $P = .03$ ) and periprosthetic fractures due to falls ( $P = .01$ ), whereas men showed a higher incidence of periprosthetic joint infections ( $P < .01$ ).

## CONCLUSION

This study found that female patients undergo SA at an older age and begin with worse shoulder range of motion and outcome scores than male patients. Although women experienced a greater improvement postoperatively in outcome scores and range of motion, this improvement did not reach the minimal clinically important difference. These findings suggest that male and female patients can expect similar improvements in function after undergoing SA; however, the incidence of complications may vary depending on sex.



# Effects of Body Mass Index on Outcomes in Total Shoulder Arthroplasty

---

**Mau EM, Roche CP, Zuckerman JD.**

**Bull Hosp Jt Dis (2013). 2015 Dec;73 Suppl 1:S99-106.**

Body Mass Index (BMI) is one of the metrics used to assess overall health and has been implicated in having predictive value in many aspects of health, including outcomes after shoulder replacement surgery. Outcome data from a multi-institutional database with an average follow-up period of 39.8 months (minimum 24-months) demonstrated that all patients, regardless of BMI, improved significantly after treatment with anatomic total shoulder arthroplasty (aTSA) or reverse total shoulder arthroplasty (rTSA). Improvements in outcomes were stratified and compared based upon BMI in three groups: less than 25, 25 to 35, and greater than 35. Comparing these measures demonstrated that aTSA patients with higher BMI were generally associated with lower functional postoperative outcome metric scores than aTSA patients with lower BMI, though the preoperative to postoperative gains were generally equivalent regardless of BMI. Interestingly, postoperative outcome metric scores with rTSA patients were equivalent regardless of BMI as were the pre-to-postoperative gains.

Additionally, differences in the magnitude of pre-to-postoperative improvement of range of motion (ROM) measurements between patients of BMI less than 25 and BMI greater than 35 were noted for forward flexion, internal rotation, and active and passive external rotation. The actual clinical significance of these differences is unknown. Finally, patients with lower BMI appeared to have a higher incidence of low-grade scapular notching.



# Using Machine Learning to Predict Internal Rotation after Anatomic and Reverse Total Shoulder Arthroplasty

**Kumar, V. et al. Using Machine Learning to Predict Internal Rotation after Anatomic and Reverse Total Shoulder Arthroplasty. JSES. In press. 2021**

## BACKGROUND

Improvement in internal rotation (IR) after anatomic (aTSA) and reverse (rTSA) total shoulder arthroplasty is difficult to predict, with rTSA patients experiencing greater variability and more limited IR improvements than aTSA patients. The purpose of this study is to quantify and compare the IR score for aTSA and rTSA patients and create supervised machine learning that predicts IR after aTSA and rTSA at multiple postoperative time points.

## METHODS

Clinical data from 2270 aTSA and 4198 rTSA patients were analyzed using 3 supervised machine learning techniques to create predictive models for internal rotation as measured by the IR score at 6 postoperative time points. Predictions were performed using the full input feature set and 2 minimal input feature sets. The mean absolute error (MAE) quantified the difference between actual and predicted IR scores for each model at each time point. The predictive accuracy of the XGBoost algorithm was also quantified by its ability to distinguish which patients would achieve clinical improvement greater than the minimal clinically important difference (MCID) and substantial clinical benefit (SCB) patient satisfaction thresholds for IR score at 2-3 years after surgery.

## RESULTS

rTSA patients had significantly lower mean IR scores and significantly less mean IR score improvement than aTSA patients at each postoperative time point. Both aTSA and rTSA patients experienced significant improvements in their ability to perform activities Background: Improvement in internal rotation (IR) after anatomic (aTSA) and reverse (rTSA) total shoulder arthroplasty is difficult to predict, with rTSA patients experiencing

greater variability and more limited IR improvements than aTSA patients. The purpose of this study is to quantify and compare the IR score for aTSA and rTSA patients and create supervised machine learning that predicts IR after aTSA and rTSA at multiple postoperative time points of daily living (ADLs); however, aTSA patients were significantly more likely to perform these ADLs. Using a minimal feature set of preoperative inputs, our machine learning algorithms had equivalent accuracy when predicting IR score for both aTSA (0.92-1.18 MAE) and rTSA (1.03-1.25 MAE) from 3 months to >5 years after surgery. Furthermore, these predictive algorithms identified with 90% accuracy for aTSA and 85% accuracy for rTSA which patients will achieve MCID IR score improvement and predicted with 85% accuracy for aTSA patients and 77% accuracy for rTSA which patients will achieve SCB IR score improvement at 2-3 years after surgery.

## DISCUSSION

Our machine learning study demonstrates that active internal rotation can be accurately predicted after aTSA and rTSA at multiple postoperative time points using a minimal feature set of preoperative inputs. These predictive algorithms accurately identified which patients will, and will not, achieve clinical improvement in IR score that exceeds the MCID and SCB patient satisfaction thresholds.

# Validation of a Machine Learning Derived Clinical Metric to Quantify Outcomes after TSA

Roche, C. et al. Validation of a Machine Learning Derived Clinical Metric to Quantify Outcomes after TSA. *J Shoulder Elbow Surg.* 2021 Feb 16;S1058-2746(21)00101-4.

## BACKGROUND

We propose a new clinical assessment tool constructed using machine learning, called the Shoulder Arthroplasty Smart (SAS) score to quantify outcomes following total shoulder arthroplasty (TSA).

## METHODS

Clinical data from 3667 TSA patients with 8104 postoperative follow-up reports were used to quantify the psychometric properties of validity, responsiveness, and clinical interpretability for the proposed SAS score and each of the Simple Shoulder Test (SST), Constant, American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form (ASES), University of California Los Angeles (UCLA), and Shoulder Pain and Disability Index (SPADI) scores.

## RESULTS

Convergent construct validity was demonstrated, with all 6 outcome measures being moderately to highly correlated preoperatively and highly correlated postoperatively when quantifying TSA outcomes. The SAS score was most correlated with the UCLA score and least correlated with the SST. No clinical outcome score exhibited significant floor effects preoperatively or postoperatively or significant ceiling effects preoperatively; however, significant ceiling effects occurred postoperatively for each of the SST (44.3%), UCLA (13.9%), ASES (18.7%), and SPADI (19.3%) measures. Ceiling effects were more pronounced for anatomic than reverse TSA, and generally, men, younger patients, and whites who received TSA were more likely to experience a ceiling effect than TSA patients who were female,

older, and of non-white race or ethnicity. The SAS score had the least number of patients with floor and ceiling effects and also exhibited no response bias in any patient characteristic analyzed in this study. Regarding clinical interpretability, patient satisfaction anchor-based thresholds for minimal clinically importance difference and substantial clinical benefit were quantified for all 6 outcome measures; the SAS score thresholds were most similar in magnitude to the Constant score. Regarding responsiveness, all 6 outcome measures detected a large effect, with the UCLA exhibiting the most responsiveness and the SST exhibiting the least. Finally, each of the SAS, ASES, Constant, and SPADI scores had similarly large standardized response mean and effect size responsiveness.

## DISCUSSION

The 6-question SAS score is an efficient TSA-specific outcome measure with equivalent or better validity, responsiveness, and clinical interpretability as 5 other historical assessment tools. The SAS score has an appropriate response range without floor or ceiling effects and without bias in any target patient characteristic, unlike the age, gender, or race/ethnicity bias observed in the ceiling scores with the other outcome measures. Because of these substantial benefits, we recommend the use of the new SAS score for quantifying TSA outcomes.

# Development of a Predictive Model for a Machine Learning Derived Shoulder Arthroplasty Clinical Outcome Score

**Kumar, V. et al. Development of a Predictive Model for a Machine Learning Derived Shoulder Arthroplasty Clinical Outcome Score. Seminars in Arthroplasty: JSES. 2021. In press.**

## BACKGROUND

We use machine learning to create predictive models from preoperative data to predict the Shoulder Arthroplasty Smart (SAS) score, the American Shoulder and Elbow Surgeons (ASES) score, and the Constant score at multiple postoperative time points and compare the accuracy of each algorithm for anatomic total shoulder arthroplasty (aTSA) and reverse total shoulder arthroplasty (rTSA).

## METHODS


Clinical data from 2270 patients who underwent aTSA and 4198 patients who underwent rTSA were analyzed using 3 supervised machine learning techniques to create predictive models for the SAS, ASES, and Constant scores at 6 different postoperative time points using a full input feature set and the 2 different minimal feature sets. Mean absolute errors (MAEs) quantified the difference between actual and predicted outcome scores for each model at each postoperative time point. The performance of each model was also quantified by its ability to predict improvement greater than the minimal clinically important difference (MCID) and the substantial clinical benefit (SCB) patient satisfaction thresholds for each outcome measure at 2-3 years after surgery.

## RESULTS

All 3 machine learning techniques were more accurate at predicting aTSA and rTSA outcomes using the SAS score (aTSA:  $\pm 7.41$  MAE; rTSA:  $\pm 7.79$  MAE), followed by the Constant score (aTSA:  $\pm 8.32$  MAE; rTSA:  $\pm 8.30$  MAE) and finally the ASES score (aTSA:  $\pm 10.86$  MAE; rTSA:  $\pm 10.60$  MAE). These prediction accuracy trends were maintained across the 3 different model input categories for each of the SAS, ASES, and Constant models at each postoperative time point. For patients who underwent aTSA, the XGBoost predictive models achieved 94%-97% accuracy in MCID with an area under the receiver operating curve (AUROC) between 0.90-0.97 and 89%-94% accuracy in SCB with an AUROC between 0.89-0.92 for the 3 clinical scores using the full feature set of inputs. For patients who underwent rTSA, the XGBoost predictive models achieved 95%-99% accuracy in MCID with an AUROC between 0.88-0.96 and 88%-92% accuracy in SCB with an AUROC between 0.81-0.89 for the 3 clinical scores using the full feature set of inputs.

## DISCUSSION

Our study demonstrated that the SAS score predictions are more accurate than the ASES and Constant predictions for multiple supervised machine learning techniques, despite requiring fewer input data for the SAS model. In addition, we predicted which patients will and will not achieve clinical improvement that exceeds the MCID and SCB thresholds for each score; this highly accurate predictive capability effectively riskstratifies patients for a variety of outcome measures using only preoperative data.



# Use of Machine Learning to Assess the Predictive Value of 3 Commonly Used Clinical Measures to Quantify Outcomes After Total Shoulder Arthroplasty

---

**Kumar, V. et al. Use of Machine Learning to Assess the Predictive Value of 3 Commonly Used Clinical Measures to Quantify Outcomes After Total Shoulder Arthroplasty. Seminars in Arthroplasty: JSES. 2021**

## BACKGROUND

An important psychometric parameter of validity that is rarely assessed is predictive value. In this study we utilize machine learning to analyze the predictive value of 3 commonly used clinical measures to assess 2-year outcomes after total shoulder arthroplasty (TSA).

## METHODS

XGBoost was used to analyze data from 2790 TSA patients and create predictive algorithms for the American Shoulder and Elbow Surgeons (ASES), Constant, and the University of California Los Angeles (UCLA) scores and also quantify the most meaningful predictive features utilized by these measures and for all questions comprising each measure to rank and compare their value to predict 2-year outcomes after TSA.

## RESULTS

Our results demonstrate that the ASES, Constant, and UCLA measures rarely considered the most-predictive features relevant to 2-year TSA outcomes and that each outcome measure was composed of questions with different distributions of predictive value. Specifically, the questions composing the UCLA score were of greater predictive value than the Constant questions, and the questions composing the Constant score were of greater predictive value than the ASES questions. We also found the preoperative Shoulder Pain and Disability Index (SPADI) score to be of greater predictive value than the preoperative ASES, Constant, and UCLA scores. Finally, we identified the types of preoperative input questions that were most-predictive (subjective self-assessments of pain and objective measurements of active range of motion and strength) and also those that were least-predictive of 2-year TSA outcomes (subjective task-specific activities of daily living questions).

## DISCUSSION

Machine learning can quantify the predictive value of the ASES, Constant, and UCLA scores after TSA. Future work should utilize this and related techniques to construct a more efficient and effective clinical outcome measure that incorporates subjective and objective input questions to better account for the preoperative factors that influence postoperative outcomes after TSA.

# Predict+ Relevant Research Publications

## References

1. **Kumar, V. et al.** What Is the Accuracy of Three Different Machine Learning Techniques to Predict Clinical Outcomes After Shoulder Arthroplasty? *Clinical Orthopedics and Related Research*. In Press. 2020.
2. **Kumar, V. et al.** Using Machine Learning to Predict Clinical Outcomes After Shoulder Arthroplasty with a Minimal Feature Set. *JSES*. In Press. 2020.
3. **Parsons, M. et al.** Preoperative External Rotation Deficit Does Not Predict Poor Outcomes or Lack of Improvement After Reverse Total Shoulder Arthroplasty. *Journal of Orthopaedics*. 2020.
4. **Parsons, M. et al.** Patient-reported Outcomes for Reverse Total Shoulder Arthroplasty – A Comparative Risk Factor Analysis of Improved Versus Unimproved Cases. *JSES Open Access*. 2019 Sep 13;3(3):174-178.
5. **Friedman, R. et al.** Preoperative Parameters That Predict Postoperative Patient Reported Outcome Measures and Range of Motion with Anatomic and Reverse Total Shoulder Arthroplasty. *JSES Open Access*. In Press. 2019.
6. **Parada, S. et al.** Comparison of Complication Types and Rates Associated with Anatomic and Reverse Total Shoulder Arthroplasty. *JSES*. 2020.
7. **Flurin, PH. et al.** Comparison of Survivorship and performance of a platform shoulder system in anatomic and reverse total shoulder arthroplasty. *JSES International*. In Press. 2020.
8. **Schoch, B. et al.** Deltoid Fatigue: A Longitudinal Assessment of Reverse Shoulder Arthroplasty Over Time. *JSES*. In press. 2020.
9. **Schoch, B. et al.** Defining the tipping point for primary shoulder arthroplasty. *JSES Open Access*. 2019. Volume 3, Issue 4, December 2019, Pages 273-277.
10. **Friedman, R. et al.** Are age and patient gender associated with different rates and magnitudes of clinical improvement after Reverse Shoulder Arthroplasty? *Clin Orthop Relat Res*. 2018 Jun;476(6):1264-1273.
11. **Simovitch, R. et al.** Rate of Improvement in Clinical Outcomes with Anatomic and Reverse Total Shoulder Arthroplasty. *J Bone Joint Surg Am*. 2017 Nov 1;99(21):1801-1811.
12. **Simovitch, R. et al.** Quantifying Success after Total Shoulder Arthroplasty: the minimal clinically important difference. *J Shoulder Elbow Surg*. 2018 Feb;27(2):298-305.
13. **Simovitch, R. et al.** Quantifying Success after Total Shoulder Arthroplasty: the Substantial Clinical Benefit. *J Shoulder Elbow Surg*. 2018 May;27(5):903-911.
14. **Eichinger, J. et al.** The Effect of BMI on Internal Rotation and Function Following Anatomic and Reverse Total Shoulder Arthroplasty. *JSES*. In press. 2020.
15. **Patel, R. et al.** Results of Total Shoulder Arthroplasty in Patients Under vs. Over Fifty-Five Years of Age: An Analysis of 1,135 Patients with over 2 Years of Follow-up. *J Shoulder Elbow Surg*. 2019.
16. **Okoroha, KR. et al.** Early Outcomes of Shoulder Arthroplasty According to Gender. *JSES Open Access*. 2019 Mar 16;3(1):43-47.
17. **Mau, E. et al.** Effects of Body Mass Index on Outcomes in Total Shoulder Arthroplasty. *Bulletin for the Hospital for Joint Diseases*. Vol. 73: S99-106. 2015.
18. **Kumar, V. et al.** Using Machine Learning to Predict Internal Rotation after Anatomic and Reverse Total Shoulder Arthroplasty. *JSES*. In press. 2021.
19. **Roche, C. et al.** Validation of a Machine Learning Derived Clinical Metric to Quantify Outcomes after TSA. *J Shoulder Elbow Surg*. 2021 Feb 16:S1058-2746(21)00101-4.
20. **Kumar, V. et al.** Development of a Predictive Model for a Machine Learning Derived Shoulder Arthroplasty Clinical Outcome Score. *Seminars in Arthroplasty: JSES*. 2021. In press.
21. **Kumar, V. et al.** Use of Machine Learning to Assess the Predictive Value of 3 Commonly Used Clinical Measures to Quantify Outcomes After Total Shoulder Arthroplasty. *Seminars in Arthroplasty: JSES*. 2021

Predict+ was developed from a dataset of primary anatomic and reverse total shoulder arthroplasty patients using the Exactech Equinox platform shoulder prosthesis where patients with revisions, humeral fractures, or hemiarthroplasty were excluded; therefore, model predictions may not be appropriate for those excluded indications or other prosthesis types or designs. Predict+ is intended for labeled indications for use of the Equinox platform total shoulder system. Please consult the instructions for use accompanying the Equinox implants.

Predict+ utilizes machine learning algorithms that learned patterns from aggregate clinical outcomes data collected from >30 different sites. This clinical data inevitably contains bias inherited from the unique circumstances of surgeons, hospitals, patients, and data collection procedures. As a result, model predictions in some cases may not be representative of the outcomes achieved by patients of different demographics, regions, or ethnicity/race than comprise the clinical outcomes database from which the algorithms are derived, and model predictions may be biased against patients too sick to safely undergo the procedure or patients whose condition is not sufficiently degenerative to have the procedure. Factors influencing each patient prediction are included in the Predict+ report for independent review, and additional details on the locked predictive algorithms and their validation and limitations are available in peer-reviewed publications at [www.exac.com/predictpluserferences](http://www.exac.com/predictpluserferences).

Each patient's needs are unique and different, and patient-specific requirements for pain relief and functional improvement may not align with established thresholds for improvement. As predictions indicate a range of anticipated outcomes, Predict+ should be used to inform treatment decision-making and shall never be misused to deny treatment.

The products discussed herein may be available under different trademarks in different countries. All copyrights, and pending and registered trademarks, are property of Exactech. This material is intended for the sole use and benefit of the Exactech sales force and physicians. It should not be redistributed, duplicated or disclosed without the express written consent of Exactech. ©2022 Exactech. 12-0001480 Rev B 1222

Exactech is proud to have offices and distributors around the globe. For more information about Exactech products available in your country, please visit [www.exac.com](http://www.exac.com)



**EXACTECH, INC.**  
2320 NW 66TH COURT  
GAINESVILLE, FL 32653 USA

+1 352.377.1140  
+1 800.EXACTECH  
+1 352.378.2617 (FAX)  
[www.exac.com](http://www.exac.com)