JB&JS OPEN ACCESS

Machine Learning Can Predict Level of Improvement in Shoulder Arthroplasty

Paul B. McLendon, MD, Kaitlyn N. Christmas, BS, CCRC, Peter Simon, PhD, Otho R. Plummer, PhD, Audrey Hunt, BS, Adil S. Ahmed, MD, Mark A. Mighell, MD, and Mark A. Frankle, MD

Investigation performed at the Foundation for Orthopaedic Research and Education (F.O.R.E.), Tampa, Florida

Background: The ability to accurately predict postoperative outcomes is of considerable interest in the field of orthopaedic surgery. Machine learning has been used as a form of predictive modeling in multiple health-care settings. The purpose of the current study was to determine whether machine learning algorithms using preoperative data can predict improvement in American Shoulder and Elbow Surgeons (ASES) scores for patients with glenohumeral osteoarthritis (OA) at a minimum of 2 years after shoulder arthroplasty.

Methods: This was a retrospective cohort study that included 472 patients (472 shoulders) diagnosed with primary glenohumeral OA (mean age, 68 years; 56% male) treated with shoulder arthroplasty (431 anatomic total shoulder arthroplasty and 41 reverse total shoulder arthroplasty). Preoperative computed tomography (CT) scans were used to classify patients on the basis of glenoid and rotator cuff morphology. Preoperative and final postoperative ASES scores were used to assess the level of improvement. Patients were separated into 3 improvement ranges of approximately equal size. Machine learning methods that related patterns of these variables to outcome ranges were employed. Three modeling approaches were compared: a model with the use of all baseline variables (Model 1), a model omitting morphological variables (Model 2), and a model omitting ASES variables (Model 3).

Results: Improvement ranges of ≤28 points (class A), 29 to 55 points (class B), and >55 points (class C) were established. Using all follow-up time intervals, Model 1 gave the most accurate predictions, with probability values of 0.94, 0.95, and 0.94 for classes A, B, and C, respectively. This was followed by Model 2 (0.93, 0.80, and 0.73) and Model 3 (0.77, 0.72, and 0.71).

Conclusions: Machine learning can accurately predict the level of improvement after shoulder arthroplasty for glenohumeral OA. This may allow physicians to improve patient satisfaction by better managing expectations. These predictions were most accurate when latent variables were combined with morphological variables, suggesting that both patients' perceptions and structural pathology are critical to optimizing outcomes in shoulder arthroplasty.

Level of Evidence: Therapeutic Level IV. See Instructions for Authors for a complete description of levels of evidence.

Patient expectations can have a strong influence on postoperative outcomes and patient satisfaction in the field of orthopaedic surgery[1-3]. Studies from the total knee[4-7], hip[7,8], and shoulder[9-11] arthroplasty literature illustrate correlations between preoperative patient expectations and postoperative outcomes. Shoulder surgeons have endeavored to quantify preoperative patient factors that may influence outcomes, such as functional assessment measures[12], soft-tissue integrity[13-15], and bone loss/wear symmetry[16-18]. The myriad of assessment measures highlights both the importance and challenge of creating accurate, comprehensive predictive models for surgeons to anticipate outcomes, counsel patients, and manage expectations.

Predictive modeling has attracted considerable interest in health care, particularly since the advent of electronic medical records containing troves of patient data[19]. Strategies for assimilating these vast data into meaningful clinical application

Disclosure: DJO Surgical provided funding for this study to the Foundation for Orthopaedic Research and Education (F.O.R.E.). DJO Surgical did not have input into the design, data collection, analysis, or manuscript preparation. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work (including relationships with DJO Surgical) (http://links.lww.com/JBJSOA/A262).

include the use of simple classifications derived from medical-expert consensus, more rigorous statistical regression, and more recently, artificial intelligence via machine learning algorithms that continuously improve through experience[20]. Interestingly, these models routinely demonstrate the highest accuracy of the above strategies, often double that of traditional linear statistics[21]. However, routine clinical application of these models is still rare.

Various machine learning methods have been used to successfully construct an array of predictions, such as new diagnoses of heart failure[22], the development of diabetes mellitus[23], the recurrence of breast cancer[24], postoperative complications following spinal deformity surgery[25], and health-care utilization in terms of the cost and complexity of care[26,27]. However, the utilization and potential benefits of predictive machine learning have, to our knowledge, yet to be explored in the area of glenohumeral osteoarthritis (OA) and shoulder arthroplasty. This bears importance since shoulder arthroplasty overall is a successful procedure for the treatment of glenohumeral OA, although not all patients achieve uniform improvement[28-30]. If various preoperative objective factors and latent variables (e.g., American Shoulder and Elbow Surgeons [ASES] scores) can be analyzed and assimilated into a predictive machine learning model, patients could potentially be stratified into tiers of anticipated postoperative improvement, thus augmenting preoperative expectation counseling and ultimately, patient satisfaction.

The primary aim of the current study was to determine whether preoperative latent variables (ASES scores), morphometric measures (bone and soft-tissue deficiency), and demographic data can be combined and analyzed via machine learning algorithms to predict a maximum level of improvement for patients with glenohumeral OA at a minimum of 2 years following shoulder arthroplasty. We hypothesized that different tiers of improvement can be established with this approach and that preoperative data can reliably predict which tier of improvement a patient will ultimately achieve.

Materials and Methods

Patient Demographics, Latent Variables, and Morphology

We retrospectively reviewed the cases of 472 patients (472 shoulders) who underwent either anatomic total shoulder arthroplasty (TSA) or reverse total shoulder arthroplasty (RSA) for a diagnosis of primary glenohumeral OA performed by a single surgeon (M.A.F.) from January 1, 2007, through December 31, 2015. All patients completed ASES forms and had computed tomography (CT) scans performed on their operative shoulder prior to surgery. Patients were instructed to receive follow-up at regular postoperative intervals (3 months, 6 months, 1 year, and annually thereafter). Three hundred patients returned for their 2-year visit (range, 21 to 30 months) as instructed, and 172 patients returned at some point after this window (range, 31 to 99 months).

Shoulders were divided by preoperative morphological features as assessed using CT scans. To allow for reformatting in the plane of the scapula, 2-dimensional (2D) CT images were imported into Mimics software (version 14.1; Materialise). The determination of glenoid morphology was based on these reformatted 2D axial images. Morphology was assigned using the original Walch classification system[31], consisting of the subtypes A1, A2, B1, B2, and C (inter- and intrarater reliability: 0.605 and 0.790, respectively). While the modified Walch classification has shown superior reliability when 3D CT reconstructions are used to analyze the scapula as a free body[32], we found that we did not have adequate numbers in many of the additional new subgroups to provide useful analysis. Fatty infiltration (FI) of the rotator cuff was assigned on the basis of the Goutallier classification system (0 = normal muscle, 1 = fatty streaks within muscle, 2 = less fat than muscle, 3 = equal amounts of fat and muscle, 4 = more fat than muscle)[15]. Atrophy of the supraspinatus was assessed using the tangent sign of Zanetti et al.[33]. The most lateral image in which the scapular spine and coracoid process were in contact with the body of the scapula was used, and a line was drawn from the most superior aspect of the scapular spine to the most superior aspect of the coracoid process. A positive tangent sign was assigned when the bulk of the supraspinatus muscle was below this line. Previous work has shown that use of the tangent sign is an accurate and reproducible method for assessing atrophy and that the presence of a positive tangent sign is highly correlated with the presence of supraspinatus FI[14]. Three fellowship-trained shoulder surgeons independently evaluated the CT scans for Walch type, Goutallier classification, and the tangent sign. Each of these preoperative morphometric observations was included in the study as an independent variable. Additional independent variables were patient age at the time of surgery, patient sex, operative side, preoperative visual analog scale (VAS) score for pain, preoperative responses to the 10 activity-specific ASES function questions, and the ASES total score (Table I).

Creating Models with Differing Follow-up Intervals

To gauge the level of improvement, the change in individual patient ASES total scores from preoperative baseline to 2-year postoperative follow-up (range, 21 to 30 months) was analyzed. Patients were separated into 3 improvement ranges of approximately equal size. We did not define 2-year follow-up as a strict 24 months. This is because many of our patients are only part-time local residents, spending the remainder of the year in other states or abroad. Therefore, patients who received follow-up within the period of 21 to 30 months were classified as those with “2-year-range” follow-up. This group of patients was included in the model, with 2-year-range follow-up as an independent variable for outcome prediction. An additional subset of patients did not show for their 2-year-range follow-up. They instead returned at later intervals, from 31 to 99 months postoperatively. For this group of patients, the ASES total score from the next-closest follow-up visit after the 2-year range was used. A separate model was created using this later follow-up interval to evaluate and compare the effect of different follow-up time points in predicting the level of improvement.

Three modeling approaches were compared: a model with the use of all baseline variables (Model 1), a model

TABLE I Summary of Patient Variables Used as Inputs for Machine Learning Models

Label	Description	Values
Explanatory variables		
Sex	Patient's sex	Male, female
Age	Patient's age	Numeric value between 28 and 89
Walch	Walch classification	A1, A2, B1, B2, C
Tangent	Tangent sign	0, 1
GoutSupra	Goutallier classification of the supraspinatus	0, 1, 2, 3, 4
GoutInfra	Goutallier classification of the infraspinatus	0, 1, 2, 3, 4
GoutTeres	Goutallier classification of the teres minor	0, 1, 2, 3, 4
Subscap	Goutallier classification of the subscapularis	0, 1, 2, 3, 4
OperativeSide	Patient's operative side	RT (right), LT (left)
ASES_PreOp	Preop. ASES total score	Continuous, 0 to 100
ASES function question number		
Q1	Put on a coat (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q2	Sleep on your painful or affected side (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q3	Wash back/do up bra in back (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q4	Manage toileting (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q5	Q6	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
	Reach a high shelf (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q7	Lift 10 lb above the shoulder (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q8	Throw a ball overhand (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q9	Do usual work (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Q10	Do usual sport (preop.)	0 (unable to do), 1 (very difficult to do), 2 (somewhat difficult), 3 (not difficult)
Pain	How bad is your pain today? (at preop.)	0 to 10
Follow-up	Months postop.	21 to 99
Target variable		
Difference	ASES 2-yr postop. score − ASES preop. score	−36.67 to 98.33 (target variable)

omitting morphological variables (Model 2), and a model omitting ASES variables (Model 3). The specific machine learning models used in the present study were created by OBERD–Universal Research Solutions (www.oberd.com) and are described in detail in Appendix I. The operative technique and rehabilitation are detailed in Appendix II.

Results

General Outcomes Data

In this cohort, 431 patients were treated with TSA and 41 patients were treated with RSA. There were 266 male (56%) and 206 female patients (44%), with an overall average age at the time of surgery of 68 years (range, 28 to 89 years). Additional age information is shown in Table II, and Figure 1 shows the age distribution density. Patient ASES scores obtained preoperatively and 2 years postoperatively, and the ultimate difference between these scores, are shown in Table III.

Predictive Ability of Machine Learning Models

The difference between the ASES scores (2-year-range postoperative ASES score minus preoperative ASES score) for each patient produced the probability density distribution shown in

TABLE II Patient Age at the Time of Surgery

	Age *(yr)*
Mean	68
Median	69
25% quartile	64
75% quartile	74
Min.	28
Max.	89

Figure 2. The model indicated that a preponderance of patients would have a clinically important improvement but that some patients will be worse at the postoperative 2-year range. The modeling can provide additional information that would be better able to predict a specific patient's outcome at the 2-year range postoperatively (see Appendix I).

Effect of Follow-up on Predictive Ability

A total of 300 patients were available for the 2-year-range follow-up (21 to 30 months). When the model was restricted to only patients with available 2-year-range follow-up, the ability to predict change in the ASES total score demonstrated good sensitivity, as seen in Table IV.

A total of 172 patients had follow-up beyond the 2-year range. When these patients (with a follow-up interval of up to 99 months postoperatively) were added to the model, the ability to predict change in the ASES total score was enhanced when compared with the restricted model that included only those with 2-year-range follow-up. This is consistent with the fact that there was a modest decrease in scores with an increasing follow-up interval ($r = -0.14$) and that the models were able to incorporate this trend because follow-up was included as variable. This is depicted in Table V.

Effect of Input Variables on Predictive Ability

For each model, the finite set of input variables (patient demographic data, latent variables, and morphological variables) from the collected preoperative data can be drawn to form a specific prediction, expressed as follows: when a patient is predicted to be in a given improvement range, what is the probability of this being the patient's actual improvement? These probabilities are summarized in Tables IV and V for the 3 different models. For example, according to Model 1 in Table IV, if a patient is predicted to be in class B, then the probability, p(A), of actually being in class A is 0.10, the probability, p(B), of actually being in class B is 0.87, and the probability, p(C), of actually being in class C is 0.03. The sensitivity is the probability that an actual member of class B is correctly predicted to be in class B (0.89 in Model 1).

For any given patient, the model can usually determine the classification without using all baseline variables. In Model 1, every variable except ASES function question number 9 was relevant for >50% of the patients. The most valuable variables were ASES total score, ASES function question number 2, Walch classification, and VAS for pain (>95% of cases), followed by operative side, subscapularis FI, ASES function question number 7, supraspinatus FI, and age (>85% of cases). For Model 2, the most valuable variables were preoperative ASES total score, age, ASES function question number 8, VAS for pain, and ASES function question number 3 (>90% of patients), followed by sex, ASES function question number 2, and ASES function question number 5 (80% of patients). For Model 3, only Walch classification, tangent sign, and subscapularis FI were relevant in >50% of the cases.

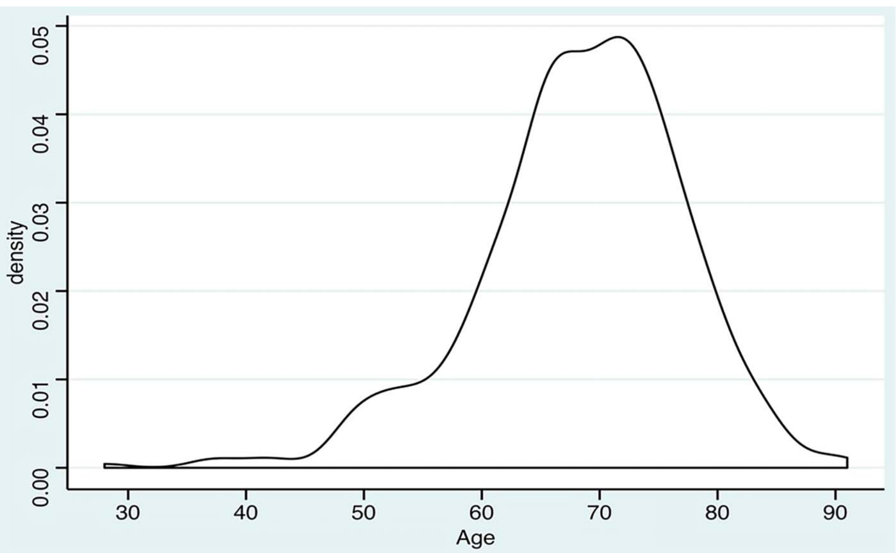

Fig. 1

Patient age distribution density. Age is given in years.

TABLE III Mean Change in ASES Scores

Measure	Mean Score All Patients (N = 472)	TSA (N = 431)	RSA (N = 41)
Preop. ASES	36.5	37.0	31.1
2-yr ASES	78.1	78.5	73.5
Difference in ASES	41.6	41.5	42.4

Discussion

This study provided evidence that preoperative data can accurately predict the level of improvement achieved at 2 years and beyond following shoulder replacement for glenohumeral OA. We used computer-based predictive methods derived from machine learning algorithms. The data available for analysis included observations regarding patients' morphological features, ASES scores, patient age, sex, and follow-up duration.

The best results were obtained when all variables were made available to the algorithms, i.e., both patient-reported and observed morphological data. The synergy between these 2 types of data is evident in Model 1 (Tables IV and V). Neither type of variable could be excluded without a loss of accuracy, as shown in Models 2 and 3; they are complementary components of outcome prediction in this patient population. The next best results were obtained with Model 2, where it was observed that the individual ASES functional questions were able to make accurate predictions. Of note, the total ASES score alone was insufficient to make accurate predictions and therefore should be combined with patient age if no other variables are available.

This study utilized machine learning to recognize various clusters of patterns in preoperative variables to predict the surgical outcomes for patients who exhibited those patterns. The predictions consisted of probabilities of the occurrence of possible outcomes for a given patient whose data match the pattern. A collection of patterns is termed a "predictive model," and the model can be used to obtain a unique prediction for any patient for whom the preoperative data are known.

A desirable characteristic of this approach is that it is self-correcting: any noninfluential variables will not end up in the model, and redundant variables will be avoided no matter their initial usefulness. There is no a priori assumption about normality (or other parametric forms) of the distribution of variables, errors, or outcomes, and it can be applied to non-numerical data such as ethnicity, sex, comorbidities, and the like. Both sampling error and flaws in the model are reflected in the results so that the overall accuracy can be assessed.

Some drawbacks of this approach are that it requires relatively large data sets of known outcomes to build an accurate model and to ascertain the realm of applicability. It also can be more sensitive to outliers or "bad" data. Nonetheless, if the data set is large and known to be representative

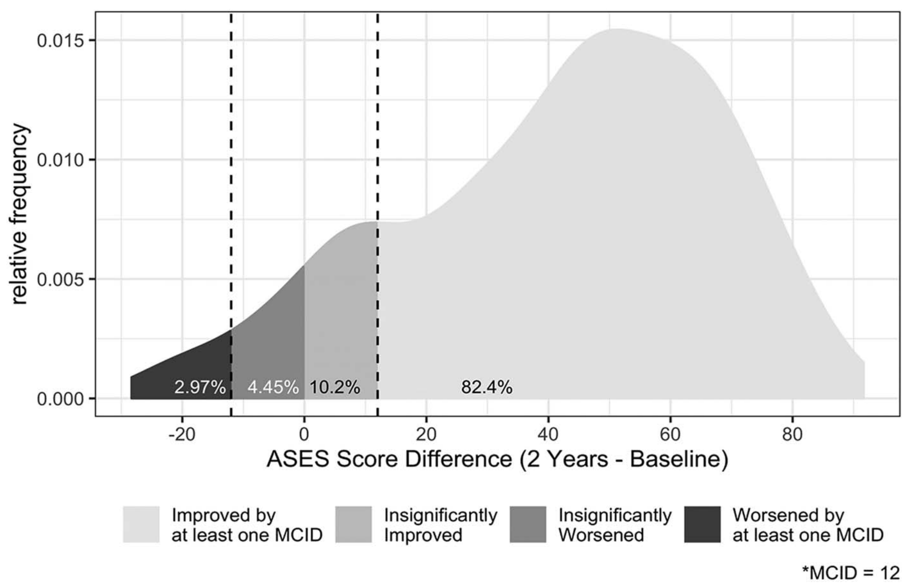

Fig. 2

Probability density distribution of the difference in ASES scores (2-year postoperative minus preoperative score). MCID = minimal clinically important difference.

TABLE IV Tier as Predicted by Different Models Using Change in ASES Score at 2-Year-Range Follow-up*

	Class†		
	A	B	C
Model 1 predicted tier			
Probability			
p(A)	0.92	0.10	0.05
p(B)	0.06	0.87	0.06
p(C)	0.02	0.03	0.89
Sensitivity	0.84	0.89	0.95
Model 2 predicted tier			
Probability			
p(A)	0.83	0.10	0.11
p(B)	0.12	0.86	0.17
p(C)	0.05	0.04	0.72
Sensitivity	0.78	0.69	0.92
Model 3 predicted tier			
Probability			
p(A)	0.69	0.17	0.21
p(B)	0.21	0.64	0.21
p(C)	0.10	0.19	0.58
Sensitivity	0.60	0.59	0.72

*Model 1 = all baseline variables used, Model 2 = morphological variables omitted, and Model 3 = ASES variables omitted. †Classes are separated by pre- to postoperative improvement in ASES total score, where Class A represents an improvement of ≤28 points, Class B represents an improvement of 29 to 55 points, and Class C represents an improvement of >55 points.

of a population of interest, then the model can be applied with confidence to new cases from the population.

Limitations of our study include the complexities of the machine learning analysis. This cannot be performed by hand and requires use of machine learning programs, which are not readily available to most orthopaedic surgeons. Additionally, this is a retrospective review, making it prone to the shortcomings and biases associated with all retrospective studies, including issues related to missing follow-up data. In a randomized controlled trial, random sampling serves to make the sample representative of the population being studied; in the present case, all relevant patients were included, and there were no limiting criteria except for the treatment and the availability of the study variables. We believe that patient characteristics and the surgical procedure have been described in sufficient detail to permit a surgeon to judge the applicability of the findings to other patients.

Another possible concern regarding the generalizability of the findings is whether the learning algorithms themselves inadvertently fitted peculiarities in the data that would limit the applicability to other data, a phenomenon known as "overfitting." We could not employ the common strategy of withholding some of the cases from the development process of the model to serve as independent data for testing the model. The complexity of the present study required that both sets be large enough to avoid skewing by outliers, and there were not enough cases to permit this. This limitation was substantially mitigated by the refinement process, which avoided conclusions that did not apply generally to the entire data set.

One of the major strengths of this study is the large volume of data used for our machine analysis. These data were obtained from a consecutive series of patients treated by a single surgeon using a well-described, standardized operative technique and rehabilitation protocol. The approximate 40-point average improvement in ASES scores after shoulder arthroplasty is comparable with that of other studies in the literature[34,35]. The categorization of patients according to preoperative morphology was based on rigorous analysis by 3 independent fellowship-trained shoulder surgeons. The percentages observed for the 3 major Walch subtypes are similar to what was noted in the initial description of the classification system by Walch et al.[31]. Additionally, the impact of follow-up duration was analyzed and, when adjusting for a difference in follow-up duration, the predictive model was able differentiate outcomes if the latent and morphological factors were utilized.

In conclusion, machine learning can accurately predict the level of improvement after shoulder arthroplasty for

TABLE V Tier as Predicted by Different Models Using Change in ASES Score Beyond 2-Year-Range Follow-up

	Class*		
	A	B	C
Model 1 predicted tier			
Probability			
p(A)	0.94	0.04	0.04
p(B)	0.05	0.95	0.03
p(C)	0.02	0.01	0.94
Sensitivity	0.91	0.94	0.98
Model 2 predicted tier			
Probability			
p(A)	0.93	0.16	0.14
p(B)	0.06	0.80	0.13
p(C)	0.01	0.03	0.73
Sensitivity	0.57	0.81	0.96
Model 3 predicted tier			
Probability			
p(A)	0.77	0.17	0.06
p(B)	0.18	0.72	0.10
p(C)	0.13	0.16	0.71
Sensitivity	0.6	0.72	0.86

*Classes are separated by pre- to postoperative improvement in ASES total score, where Class A represents an improvement of ≤28 points, Class B represents an improvement of 29 to 55 points, and Class C represents an improvement of >55 points.

glenohumeral OA. This may allow physicians to improve patient satisfaction by better managing expectations. In the current study, we found that predictions were most accurate when latent variables were combined with morphological variables, suggesting that both patients' perceptions and structural pathology are critical to optimizing outcomes.

Appendix

eA Supporting material provided by the authors is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSOA/A263). ■

Note: The authors thank Emilie Song for her tireless efforts in the procurement and sorting of data for this project.

Paul B. McLendon, MD[1]
Kaitlyn N. Christmas, BS, CCRC[2]
Peter Simon, PhD[2]
Otho R. Plummer, PhD[3]
Audrey Hunt, BS[3]
Adil S. Ahmed, MD[4]
Mark A. Mighell, MD[1]
Mark A. Frankle, MD[1]

[1]Shoulder and Elbow Service, Florida Orthopaedic Institute, Tampa, Florida

[2]Translational Research, Foundation for Orthopaedic Research and Education (F.O.R.E.), Tampa, Florida

[3]Universal Research Solutions, LLC, Columbia, Missouri

[4]Department of Orthopaedic Surgery, University of South Florida Morsani College of Medicine, Tampa, Florida

Email address for M.A. Frankle: mfrankle@floridaortho.com

ORCID iD for P.B. McLendon: 0000-0002-3981-4250
ORCID iD for K.N. Christmas: 0000-0002-2220-9185
ORCID iD for P. Simon: 0000-0001-8357-0268
ORCID iD for O.R. Plummer: 0000-0002-6037-4069
ORCID iD for A. Hunt: 0000-0001-5197-9413
ORCID iD for A.S. Ahmed: 0000-0002-7064-8430
ORCID iD for M.A. Mighell: 0000-0001-8044-9411
ORCID iD for M.A. Frankle: 0000-0001-6427-0665

References

1. Gagnier JJ. Patient reported outcomes in orthopaedics. J Orthop Res. 2017 Oct;35(10):2098-108. Epub 2017 Jun 13.
2. Graham B, Green A, James M, Katz J, Swiontkowski M. Measuring patient satisfaction in orthopaedic surgery. J Bone Joint Surg Am. 2015 Jan 7;97(1):80-4.
3. Zywiel MG, Mahomed A, Gandhi R, Perruccio AV, Mahomed NN. Measuring expectations in orthopaedic surgery: a systematic review. Clin Orthop Relat Res. 2013 Nov;471(11):3446-56.
4. Deakin AH, Smith MA, Wallace DT, Smith EJ, Sarungi M. Fulfilment of preoperative expectations and postoperative patient satisfaction after total knee replacement. A prospective analysis of 200 patients. Knee. 2019 Dec;26(6):1403-12. Epub 2019 Aug 29.
5. Jain D, Nguyen LL, Bendich I, Nguyen LL, Lewis CG, Huddleston JI, Duwelius PJ, Feeley BT, Bozic KJ. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. J Arthroplasty. 2017 Sep;32(9S):S166-70. Epub 2017 Jan 18.
6. Culliton SE, Bryant DM, Overend TJ, MacDonald SJ, Chesworth BM. The relationship between expectations and satisfaction in patients undergoing primary total knee arthroplasty. J Arthroplasty. 2012 Mar;27(3):490-2. Epub 2011 Nov 23.
7. Neuprez A, Delcour JP, Fatemi F, Gillet P, Crielaard JM, Bruyère O, Reginster JY. Patients' expectations impact their satisfaction following total hip or knee arthroplasty. PLoS One. 2016 Dec 15;11(12):e0167911.
8. Jain D, Bendich I, Nguyen LL, Nguyen LL, Lewis CG, Huddleston JI, Duwelius PJ, Feeley BT, Bozic KJ. Do patient expectations influence patient-reported outcomes and satisfaction in total hip arthroplasty? A prospective, multicenter study. J Arthroplasty. 2017 Nov;32(11):3322-7. Epub 2017 Jun 16.
9. Booker S, Alfahad N, Scott M, Gooding B, Wallace WA. Use of scoring systems for assessing and reporting the outcome results from shoulder surgery and arthroplasty. World J Orthop. 2015 Mar 18;6(2):244-51.
10. Swarup I, Henn CM, Nguyen JT, Dines DM, Craig EV, Warren RF, Gulotta LV, Henn RF III. Effect of pre-operative expectations on the outcomes following total shoulder arthroplasty. Bone Joint J. 2017 Sep;99-B(9):1190-6.
11. Rauck RC, Swarup I, Chang B, Dines DM, Warren RF, Gulotta LV, Henn RF 3rd. Effect of preoperative patient expectations on outcomes after reverse total shoulder arthroplasty. J Shoulder Elbow Surg. 2018 Nov;27(11):e323-9. Epub 2018 Jun 28.
12. Lapner PL, Jiang L, Zhang T, Athwal GS. Rotator cuff fatty infiltration and atrophy are associated with functional outcomes in anatomic shoulder arthroplasty. Clin Orthop Relat Res. 2015 Feb;473(2):674-82. Epub 2014 Sep 30.
13. Rulewicz GJ, Beaty S, Hawkins RJ, Kissenberth MJ. Supraspinatus atrophy as a predictor of rotator cuff tear size: an MRI study utilizing the tangent sign. J Shoulder Elbow Surg. 2013 Jun;22(6):e6-10. Epub 2013 Jan 23.
14. Williams MD, Lädermann A, Melis B, Barthelemy R, Walch G. Fatty infiltration of the supraspinatus: a reliability study. J Shoulder Elbow Surg. 2009 Jul-Aug;18(4):581-7.
15. Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. Clin Orthop Relat Res. 1994 Jul;304:78-83.
16. Shapiro TA, McGarry MH, Gupta R, Lee YS, Lee TQ. Biomechanical effects of glenoid retroversion in total shoulder arthroplasty. J Shoulder Elbow Surg. 2007 May-Jun;16(3)(Suppl):S90-5. Epub 2006 Dec 12.
17. Mansat P, Briot J, Mansat M, Swider P. Evaluation of the glenoid implant survival using a biomechanical finite element analysis: influence of the implant design, bone properties, and loading location. J Shoulder Elbow Surg. 2007 May-Jun;16(3)(Suppl):S79-83. Epub 2006 Aug 7.
18. Luedke C, Kissenberth MJ, Tolan SJ, Hawkins RJ, Tokish JM. Outcomes of anatomic total shoulder arthroplasty with B2 glenoids: a systematic review. JBJS Rev. 2018 Apr;6(4):e7.
19. Koh HC, Tan G. Data mining applications in healthcare. J Healthc Inf Manag. 2005 Spring;19(2):64-72.
20. Neeman T. Clinical prediction models: a practical approach to development, validation, and updating by Ewout W. Steyerberg. International Statistical Review. 2009 Aug;77(2):320-1.
21. Kuhn M, Johnson K. Applied predictive modeling. Springer; 2013.
22. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inform Assoc. 2017 Mar 1;24(2):361-70.
23. Luo G. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. Health Inf Sci Syst. 2016 Mar 8;4:2.
24. Ahmad LG, Eshlaghy AT, Poorebahimi A, Ebrahimi M, Razavi AR. Using three machine learning techniques for predicting breast cancer recurrence. J Health Med Inform. 2013;4(2):1-3.
25. Kim JS, Arvind V, Oermann EK, Kaji D, Ranson W, Ukogu C, Hussain AK, Caridi J, Cho SK. Predicting surgical complications in patients undergoing elective adult spinal deformity procedures using machine learning. Spine Deform. 2018 Nov - Dec;6(6):762-70.
26. Yang C, Delcher C, Shenkman E, Ranka S. Machine learning approaches for predicting high cost high need patient expenditures in health care. Biomed Eng Online. 2018 Nov 20;17(Suppl 1):131.
27. Chen S, Bergman D, Miller K, Kavanagh A, Frownfelter J, Showalter J. Using applied machine learning to predict healthcare utilization based on socioeconomic determinants of care. Am J Manag Care. 2020 Jan;26(1):26-31.
28. Edwards TB, Kadakia NR, Boulahia A, Kempf JF, Boileau P, Némoz C, Walch G. A comparison of hemiarthroplasty and total shoulder arthroplasty in the treatment of primary glenohumeral osteoarthritis: results of a multicenter study. J Shoulder Elbow Surg. 2003 May-Jun;12(3):207-13.

29. Izquierdo R, Voloshin I, Edwards S, Freehill MQ, Stanwood W, Wiater JM, Watters WC 3rd, Goldberg MJ, Keith M, Turkelson CM, Wies JL, Anderson S, Boyer K, Raymond L, Sluka P; American Academy of Orthopedic Surgeons. Treatment of glenohumeral osteoarthritis. J Am Acad Orthop Surg. 2010 Jun;18(6):375-82.

30. Bartelt R, Sperling JW, Schleck CD, Cofield RH. Shoulder arthroplasty in patients aged fifty-five years or younger with osteoarthritis. J Shoulder Elbow Surg. 2011 Jan; 20(1):123-30. Epub 2010 Aug 25.

31. Walch G, Badet R, Boulahia A, Khoury A. Morphologic study of the glenoid in primary glenohumeral osteoarthritis. J Arthroplasty. 1999 Sep;14(6):756-60.

32. Bercik MJ, Kruse K 2nd, Yalizis M, Gauci MO, Chaoui J, Walch G. A modification to the Walch classification of the glenoid in primary glenohumeral osteoarthritis using three-dimensional imaging. J Shoulder Elbow Surg. 2016 Oct;25(10):1601-6. Epub 2016 Jun 6.

33. Zanetti M, Gerber C, Hodler J. Quantitative assessment of the muscles of the rotator cuff with magnetic resonance imaging. Invest Radiol.1998;33(3):163-70.

34. Hussey MM, Steen BM, Cusick MC, Cox JL, Marberry ST, Simon P, Cottrell BJ, Santoni BG, Frankle MA. The effects of glenoid wear patterns on patients with osteoarthritis in total shoulder arthroplasty: an assessment of outcomes and value. J Shoulder Elbow Surg. 2015 May;24(5):682-90. Epub 2014 Dec 3.

35. Cuff D, Pupello D, Virani N, Levy J, Frankle M. Reverse shoulder arthroplasty for the treatment of rotator cuff deficiency. J Bone Joint Surg Am. 2008 Jun;90(6):1244-51.