Modern Applications of Machine Learning in Shoulder Arthroplasty

A Review

Akshar V. Patel, BS

Andrew J. Stevens, BS

Noah Mallory, BS

David Gibbs, BS

Mustaqueem Pallumeera, BS

Erryk Katayama, BS

Gregory L. Cvetanovich, MD

Julie Y. Bishop, MD

Ryan C. Rauck, MD

Investigation performed at The Ohio University Wexner Medical Center, Columbus, Ohio

Abstract

» There is increased integration of machine learning (ML) to aid clinical decision-making in orthopaedic surgery.

» ML has the ability to predict both clinical outcomes such as range of motion and complications in total shoulder arthroplasty patients.

» An increased area of focus is the ability for ML to identify implants to aid in revision surgery planning.

» In this article, we review the current applications of ML in shoulder arthroplasty and discuss future areas where it may enhance orthopaedic practice.

Artificial intelligence (AI) is an umbrella term for a set of technologies which enable computers to perform complex cognitive tasks, such as natural language processing, computer vision, or machine learning (ML)[1]. ML is a groundbreaking subset of AI which uses computational algorithms applied to data sets to "train" or identify relationships between certain combinations of variables which can be used to predict future outcomes[2,3]. This powerful technology is being leveraged in medicine to guide clinical activities and decision-making[4].

There are 3 classes of ML: supervised, unsupervised, and reinforcement learning. Supervised learning builds a function from prelabeled training data sets (that is, variable X and variable Y are presented together as examples to learn from), which provides structure to the learning. Once trained, the function can be applied to new unlabeled data, assuming that the same associations are true[5]. Supervised learning is the most common ML technique used in health care, typically used to predict or estimate outcomes[6-8]. Unsupervised learning, on the other hand, generates a function without a labeled data set (the association between variable X and variable Y is not already known). Instead, the machine independently finds patterns in the data, typically through algorithmic clustering, and thus, is typically used to search for unknown associations[5,6]. Reinforcement learning uses trial-and-error to generate a relationship that best achieves a certain goal[9]. For example, a machine can simulate a series of intervention options available during medical decision-making to select the best treatment plan based on individual prognostic factors[10,11]. ML has been further developed to include deep learning, which elaborates on previously described ML by incorporating multiple hidden layers of nonlinear processing through artificial neural networks (ANNs). These deep ANNs are capable of adjusting the weighting of these nonlinear functions and integrating input data at each hierarchical processing level, which ultimately improves the efficiency of learning for massive data sets[12-14].

Orthopaedic surgery is no outlier, with many possible applications of ML[7,15,16]. For

Disclosure: The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSREV/A985).

example, ML has already been applied to lower extremity arthroplasty. Algorithms can predict primary outcomes, complications, blood transfusions, opioid use, length of stay, and payment in total hip and total knee arthroplasty[17-23]. Shoulder arthroplasty is a common procedure indicated for treatment of numerous shoulder pathologies, including glenohumeral arthropathies, proximal humerus fracture, and rotator cuff arthropathy[24,25]. Success of total shoulder arthroplasty (TSA) is measured by restoration of proper function (i.e. range of motion and strength) while improving pain and limiting complications[24,26]. ML can support in planning and counseling because of its ability to predict postoperative outcomes (such as range of motion, satisfaction, expenditures, or complications) based on preoperative and perioperative measures, such as demographic characteristics, specific diagnosis, procedure type, or comorbidities[27-30]. For instance, Simmons et al. demonstrated the utility of software algorithms for guiding medical decision-making and increasing surgeon confidence in selecting the appropriate TSA approach (anatomic vs. reverse) for certain patients[31]. Simmons et al. found that the ML-based "Clinical Decision Support Tool" improved surgeon confidence regardless of prosthesis choice and surgeon experience[31]. An increased understanding and stronger predictive model of risk factors because of advancements in ML may allow physicians better decision-making and preparedness to overcome potential hurdles in the care of TSA patients.

The goal of this article was to summarize the current literature regarding the state of ML in shoulder arthroplasty. In this article, we discuss current ML as a predictor of primary outcomes, adverse outcomes, and costs and consider future applications of ML within the field.

Predicting Outcomes after Shoulder Arthroplasty

Predicting Range of Motion

The ability of shoulder arthroplasty to improve range of motion (ROM) is well documented, but the degree of improvement often varies between individual patients and types of implants and has long been difficult to predict[25,32,33]. Although numerous studies have identified preoperative factors and patient characteristics that contribute to greater improvements in patient ROM, the analyses are difficult to apply on an individual patient level to allow for accurate predictions of postoperative improvement[34]. ML techniques have shown potential to provide individualized predictions for patient ROM changes after surgery.

Kumar et al. examined 3 different algorithms for predicting internal rotation (IR) scores after anatomic TSA (aTSA) and reverse TSA (rTSA)[35]. The mean absolute error (MAE) was used to evaluate the algorithms' accuracy, which is the difference between the predicted values and the actual values, with a MAE of 0 indicating that the model was completely accurate. From a sample of 2, 270 aTSA and 4,198 rTSA patients, these algorithms were able to predict active IR scores with 0.92 to 1.18 MAE for aTSA patients and 1.03 to 1.25 MAE for rTSA patients. Two metrics (minimal clinically important difference [MCID]; substantial clinical benefit [SCB]) were also used to assess the algorithms' accuracy.

When evaluated using IR, the algorithms were able to accurately predict which patients would achieve the MCID 90% of the time with aTSA and 85% with rTSA. When the SCB IR was assessed, the model had an accuracy of 85% with aTSA and 77% with rTSA.

These results demonstrate that ML can be used to reliably predict active IR after shoulder arthroplasty.

Kumar et al. used 3 ML techniques, such as linear regression, XGBoost, and Wide and Deep, to test predictive models for active abduction, forward flexion, and external rotation[36]. The learning ability of each model was assessed by comparing them with a baseline average analysis of the patients to serve as a study control. They found that all 3 ML techniques exhibited lower MAE compared with the baseline control, with the Wide and Deep model achieving the smallest MAE, predicting active abduction to ±18 to 21°, forward elevation to ±15 to 17°, and external rotation to ±10 to 12° of actual postoperative measurements.

Taken together, these 2 studies showcase the potential of ML techniques to predict patient ROM improvements after TSA, the implications of which are quite promising for physicians and patients alike. However, an important limitation to recognize is that these studies have been completed with data from 1 implant manufacturer, Exactech. It will be critical for more surgeons to use different prostheses and analyze these data using ML to improve the accuracy of ML to predict ROM after TSA. The capability to delineate which patients will achieve meaningful improvements after surgery would aid in the physician's ability to determine suitable candidates for TSA and those that might benefit from nonoperative treatment. In addition, these predictions could help to align physician and patient expectations on what improvements are realistically achievable after surgery.

Even with an accurate algorithm, there is a possibility that patients may experience worse postoperative outcomes than predicted regarding ROM, satisfaction, and patient-reported outcomes (PROs). To properly manage patient expectations and avoid the potential risk for medicolegal litigation, it is crucial that surgeons help patients understand the nature of ML models. One approach could be for providers to explain that the preoperative predictions are based on historical trends using large patient databases but that individual results may vary because every patient is different. The providers should be transparent about the likelihood of success in each candidate and share the latest data about how accurate the predictions have been. In addition, a disclosure statement should be included in their preoperative paperwork stating that patients are aware that the values provided are predictions and may not represent the outcome for all patients.

Predicting Patient Satisfaction

Although TSA has been shown to effectively improve function and decrease pain, some studies suggest that up to 12% of patients are still unsatisfied with their postoperative results[37,38]. Several studies to date have identified factors predictive of patient dissatisfaction, but they do not provide a means to determine risk assessment at an individual patient level[39-42]. These limitations make patient satisfaction scores another important metric that algorithms have shown the potential to predict.

Polce et al. created predictive models using 5 ML algorithms to predict patient satisfaction after TSA[28]. Satisfaction was determined based on patient response to the following binary question 2 years postoperatively: "Taking into account all activities you have done during your daily life, your level of pain, and also your functional impairment, do you consider that your current state is satisfactory?"[28] A set of 10 routinely collected preoperative variables were documented and used for the predictive model, including patient demographics, primary diagnosis, preoperative exercise, and baseline Single Assessment Numerical Evaluation (SANE) score. The SANE score is used to assess patient-reported shoulder function on a scale of 0 to 100. The authors applied their technique to a set of 413 patients and were able to predict patient satisfaction with excellent discrimination (area under the curve [AUC] = 0.80) and performance (Brier score = 0.11). The support vector machine algorithm (SVM) was the best relative predictive model of the 5 algorithms used. It was able to accurately predict postoperative patient satisfaction based on preoperative SANE score, exercise and activity level, indication for surgery, insurance status, and length of time from onset of symptoms to surgery. This study illustrates a potential role for ML models to aid clinicians in determining whether patients will be satisfied with their results postoperatively.

Predicting PROs

PROs play an increasingly important role in shoulder surgery, enabling physicians to measure subjective aspects of a patient's condition that cannot be assessed through traditional means, such as physical examination or radiographic imaging[43].

McLendon et al. analyzed the ability of algorithms to predict improvement in American Shoulder and Elbow Surgeons (ASES) scores for 472 TSA patients at a minimum of 2 years of follow-up[44]. Each patient was grouped into 1 of 3 classes based on postoperative ASES score improvement. Class A defined as improvement of 28 points or less, class B defined as improvement of 29 to 55 points, and class C defined as improvement greater than 55 points. Improvement ranges of each class were established such that there would be an approximately equal number of patients in each class. Algorithms were used to determine the probability that they would accurately predict the improvement of the patients. The authors calculated the probability that their model would accurately predict ASES improvement to be 0.94, 0.95, and 0.94 for patients in classes A, B, and C, respectively. This study demonstrates that using ML techniques, preoperative data can be used to accurately predict the level of PRO improvement after TSA.

A previously mentioned study by Kumar et al.[36] used ML models to predict University of California-Los Angeles (UCLA), Constant, and VAS pain scores after TSA, quantifying their accuracy by calculating the MAE for each prediction. The researchers showed that their models were able to predict postoperative UCLA score to ±2.5 to 3.4, the Constant score to ±7.3 to 7.9, and the VAS pain score to ±1.2 to 1.4.

Clinically, these predictive models would be beneficial for determining which surgery an individual patient was a suitable candidate for or if they would fare better with nonsurgical management. A surgeon might use the model by McLendon et al. to predict that a given patient would achieve less than a 28-point improvement in their ASES score. As previous research has approximated the MCID for the ASES score to be 20.9 after TSA, the surgeon could recognize that it may be worth considering other options for this patient[45].

Predictions Using Smaller Inputs

In certain health care systems, data may not be available to the extent needed to run a large input algorithm. In addition, from a surgeon's viewpoint, the burden of gathering and computing such information may render the model unfeasible. Kumar et al. used the XGBoost ML technique to compare prediction accuracy of 2 predictive models, 1 with 291 parameters and 1 with 19 parameters[29]. The authors found that the 2 models had similar MAEs across all time points analyzed, and both performed better than the baseline control. Kumar et al. did not find a significant difference in the accuracy when comparing the 19 parameter and 291 parameter model[29]. These results exhibit the promising feasibility of using ML techniques during surgical consultations to enhance decision-making.

Predicting Adverse Outcomes, Cost, and Nonhome Discharges After Shoulder Arthroplasty

Predicting Complications

ML has been used to predict the postoperative complications that may arise subsequent to TSA procedures. Predicting these complications can allow for improved patient selection, risk stratification, and preoperative planning[46].

Devana et al. tested several different ML models for the prediction of postoperative complications and unplanned remissions. This study focused on identifying which preoperative variables were most predictive in identifying patients with postoperative complications[47]. For example, Devana et al. observed that patients with a history of implant complications are predisposed to having more implant complications after rTSA. They surmised that these patients may inherently be prone to prosthesis-related complications. In addition, Devana et al. found that teaching hospitals and protein-calorie malnutrition were the second and third most predictive factor of complications or readmissions after rTSA. However, the authors suggested that teaching hospitals may be a co-founding factor because they tend to treat more complex shoulder pathologies. Devana et al. found that the models created have the potential for improving perioperative decision-making and the informed consent process[48].

Gowd et al. tested similar ML models and concluded that ML can accurately predict postoperative complications from a random sample of a national cohort[30]. These ML results could have implications with bundled payment or reimbursements if a surgeon is able to predict which patients are at higher risk of complications, readmissions, or nonhome discharges.

Predicting Nonhome Discharges

Arvind et al. examined 5 ML algorithms to determine which risk factors contribute most to risk prediction in a series of 9,043 TSA patients[49]. The 5 ML algorithms used were the SVM, logistic regression, random forest (RF), an adaptive boosting algorithm, and neural network. The models used variables which included demographic and comorbidity data. The study found that, of the ML algorithms, SVM performed the worst while the RF classifier performed the best. SVM may have performed poorly due to its requirement for large amounts of data to train, which can be a limitation when investigating patients with low readmission and complication rates. The RF classifier performed well due its ability to specifically identify patients at both an increased and decreased risk for readmission. This study concludes by stating that ML is able to predict unplanned readmission accurately after TSA using data from a national database. The findings from this study further support the utility of ML in predicting postoperative complications, such as unplanned readmission.

The studies discussed have several limitations and areas of further development. The ML tools used in these studies may require external validation in their efficacy and utility in the clinical setting before making definitive conclusions[46]. The use of large databases for the studies discussed also serve as a limitation because of their lack of specificity for variables that may be relevant to shoulder arthroplasty[49]. Further studies must be conducted with different data sets to ascertain the external validity of the ML models discussed[46].

Predicting Costs

Gowd et al. examined the prediction of total health care cost after TSA using ML[50]. The Nationwide Readmissions Database was queried for all primary aTSA and rTSA by *International Classification of Diseases*, *10th Revision*, procedure codes and categorized by diagnoses. Costs were calculated by using the total hospital charge and each hospitals cost-to-charge ratio. Hospital characteristics, such as procedural volume stratified by calendar year and wage index, which is a measure of the mean hospital wage in a specific area, were included. The costs of unplanned readmissions were added to the total cost of admission allowing ML algorithms to predict immediate perioperative costs. A total of 49,354 patients were selected from the database with an index TSA averaging $18,843 US dollars, with readmission costs averaging $13,871. The authors found the average total cost of a TSA to be $20,567 ± $12,316 when readmission costs were also factored in. Wage index, hospital volume, patient age, readmissions, and diagnosis-related group severity were the factors which most correlated with the cost of care. Both the logistical regression and RF algorithms were equivalent in predicting the total cost of care, with an AUC of 0.83. The study concludes by stating that readmissions largely affect the variability in cumulative costs of hospital visits. When hospital characteristics (geographical area and volume) are considered, ML algorithms may predict cases with a high likelihood of increased cost and readmission. A similar study by Karnuta et al. supports these findings with an ANN algorithm which was trained to predict length of stay, discharge deposition, and inpatient charges for TSA procedures. They concluded that their ANN demonstrated a fair to good accuracy in predicting inpatient costs[51]. The results of these studies show that ML can be an effective tool in analyzing and predicting the costs of TSA in a perioperative context.

The limitations of the current implementations of ML lie in the retrospective nature of the investigations discussed and their use of a national database[50]. This suggests that demographic characteristics which physicians share in common cannot be controlled. Variables, such as operative technique, indication for surgery, and implant type, are not recorded on the database and may affect the cost of procedures. The database only registers same state readmission, which means data from patients readmitted in a different state would not be available. ML models are a product of the data inputted, so limited data pools may lead to less representative results[51].

Predicting Nonhome Discharges

ML algorithms may also be used to predict nonhome discharges. Predicting nonhome discharges can potentially decrease the postoperative length of stay at the hospital, thus reducing the risk of complications and health care costs[46].

Lopez et al. examined the usage of ML methods to predict nonhome discharges after elective TSA[46]. Andrews et al., as cited by Lopez et al., found that patients with longer hospital stays were more likely to face postoperative complications[52]. They found that the risk of adverse complications increases by nearly 6% for each additional day admitted in the hospital. Therefore, accurately predicting and preparing for nonhome discharges can both decrease the length of hospital admittance and predict and reduce the risk of patient complications. Lopez et al. identified 21, 544 cases of elective TSA procedure which met the inclusion criteria for this study. A multivariate logistical regression was used to identify variables associated with an increased risk of nonhome discharge. Risk factors included demographic statuses and comorbidities. A boosted decision tree model and ANN were then developed and tested in their ability to predict nonhome discharges and 30-day postoperative complications when supplied the variables identified. This study found that the predictive capacity, measured by the AUC, was fair for the boosted decision tree (0.788) and was good for the ANN model (0.851). Both models reported similar accuracy, with 90.3% for the boosted decision tree and 89.9% for the ANN model. For discriminate ability, the ANN model outperformed the boosted decision tree because of its multiple neural layers being able to process large amounts of data and share information with weighted connections. The findings from this study illustrate the value of ML in predicting the postoperative complications of TSA procedures in the context of nonhome discharge length.

Implant Identification

In an individual in need of a revision, the importance of correctly identifying a previously implanted shoulder for a surgeon for surgical planning can save them substantial time and cost[53]. However, this information is sometimes unavailable for various reasons. One study estimated that implant identification averages approximately 20 minutes, and as many as 10% of arthroplasty implants are not identified before surgery, highlighting the need for an accurate method for implant identification[54].

Orthopaedics has used ML to identify fractures, classify hip osteoarthritis, and determine bone age based on X-ray images[55-57]. While using ML to identify implants is not a new concept, historically, hips and knees had a much higher rate of successful identification compared with shoulder implants[53]. Some studies state the reason for this being the wide range over which an anteroposterior shoulder X-ray can be spread[53]. However, many recent studies have reported on ML algorithms that demonstrated accuracy over 90%[53,54,58]. Geng et al. conducted a promising proof of concept showing a 93% accuracy and a 0.1 second average identification time[59].

Most ML-based radiographic studies use some form of convolutional neural network (CNN), a model that excels at complex image analysis. The CNN learns by seeing hundreds of images, detecting patterns such as edges and corners, and then associating these patterns to the implant name. To prepare the model to train, a repository consisting of several hundred shoulder radiographs with identified implants of high quality and often at multiple angles is required. The algorithm then proceeds to work through the collection of radiographs multiple times, making incremental improvements with each iteration to improve its accuracy. Finally, the algorithm is exposed to a brand new set of implant images to assess the accuracy of the model. The end result is an algorithm to increase surgeons' productivity and decrease complications from previously unidentifiable implants.

Future Directions: 3D Preoperative Planning and ML

The advent of 3D computed tomography (CT) scans has led to improvements in preoperative planning. In a recent study, Reid et al. found no significant differences in errors when comparing 2D and 3D CT scan measurements, but surgeons identified higher rates of glenoid retroversion on 3D scans[60]. In addition, Lilley et al. found a high level of concordance between preoperative 3D CT scan–based planning and intraoperative decision-making. In a series of 50 TSA cases[61], Werner et al. found that surgeons changed their decision to perform an aTSA or rTSA in 7/50 (14%) patients after using a 3D CT scan[62]. While a small study, a change in arthroplasty for 14% of patients after is noteworthy. The change in decision was primarily based on the amount of glenoid retroversion and glenoid inclination calculated in a 3D scan. This is 1 area where ML can be integrated to support clinical decision-making. In a future study, a team could use an existing data set of shoulder arthroplasty patients with 2D and 3D CT scans and calculate glenoid retroversion and glenoid inclination using ML. It would be valuable to assess whether the patients who had arthroplasty failure had glenoid measurements that would have indicated them for a type of arthroplasty different from the one they received.

Limitations

There are inherent limitations with the use of ML in TSA. The accuracy of any algorithm is contingent on training preexisting data. Any incompleteness or inaccuracies in the currently available data could lead to inaccurate results. In a clinical context, this could mean suboptimal predictions about postoperative outcomes or complication rates. Currently, there is not a significant amount of literature that investigates the role of ML in shoulder arthroplasty. Thus, the application of ML in TSA may evolve as more studies are published and greater clarity is achieved about its strengths and limitations in supporting a modern orthopaedic surgeon.

Conclusion

The utility of ML in TSA will continue to rise as more physicians become comfortable with an AI-based approach to orthopaedic care. It has a strong potential to supplement but not replace clinical intuition in the realm of predicting both clinical and adverse outcomes after shoulder arthroplasty. In addition, it can support preoperative planning for revision surgeries by correctly identifying a prosthesis. ML can be an important adjunct to the modern orthopaedic surgeons' expertise, clinical acumen, and decision-making in TSA.

Source of Funding

There was no funding for this study.

Akshar V. Patel, BS[1],
Andrew J. Stevens, BS[1],
Noah Mallory, BS[1],
David Gibbs, BS[1],
Mustaqueem Pallumeera, BS[1],
Erryk Katayama, BS[1],
Gregory L. Cvetanovich, MD[1],
Julie Y. Bishop, MD[1],
Ryan C. Rauck, MD[1]

[1]Department of Orthopaedics, The Ohio State University College of Medicine, Columbus, Ohio

Email for corresponding author:
ryan.rauck@osumc.edu

References

1. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg. 2018;268(1):70-6.

2. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019;19(1):64.

3. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. JAMA. 2016;315(6):551-2.

4. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019; 380(14):1347-58.

5. Ayodele TO. Types of machine learning algorithms. New Adv Mach Learn. 2010;3:19-48.

6. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: a literature review. Front Bioeng Biotechnol. 2018;6:75.

7. Pruneski JA, Pareek A, Kunze KN, Martin RK, Karlsson J, Oeding JF, Kiapour AM, Nwachukwu BU, Williams RJ. Supervised machine learning and associated algorithms: applications in orthopedic surgery. Knee Surg Sports Traumatol Arthrosc. 2022;31(4):1196-202.

8. Kotti M, Duffell LD, Faisal AA, McGregor AH. Detecting knee osteoarthritis and its discriminating parameters using random forests. Med Eng Phys. 2017;43:19-29.

9. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. Cambridge, MA: MIT Press; 2018.

10. Zhao Y, Zeng D, Socinski MA, Kosorok MR. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. Biometrics. 2011;67(4):1422-33.

11. Jonsson A. Deep reinforcement learning in medicine. Kidney Dis (Basel). 2019;5(1):18-22.

12. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Markets. 2021;31(3):685-95.

13. Chahal A, Gulia P. Machine learning and deep learning. Int J Innov Technol Explor Eng. 2019;8(12):4910-4.

14. Dargan S, Kumar M, Ayyagari MR, Kumar G. A survey of deep learning and its applications: a new paradigm to machine learning. Arch Comput Methods Eng. 2020;27(4):1071-92.

15. Ogink PT, Groot OQ, Karhade AV, Bongers MER, Oner FC, Verlaan JJ, Schwab JH. Wide range of applications for machine-learning prediction models in orthopedic surgical outcome: a systematic review. Acta Orthop. 2021; 92(5):526-31.

16. Ramkumar PN, Haeberle HS, Bloomfield MR, Schaffer JL, Kamath AF, Patterson BM, Krebs VE. Artificial intelligence and arthroplasty at a single institution: real-world applications of machine learning to big data, value-based care, mobile health, and remote patient monitoring. J Arthroplasty. 2019;34(10):2204-9.

17. Fontana MA, Lyman S, Sarker GK, Padgett DE, MacLean CH. Can machine learning algorithms predict which patients will achieve minimally clinically important differences from total joint arthroplasty? Clin Orthop Relat Res. 2019;477(6):1267-79.

18. Haeberle HS, Helm JM, Navarro SM, Karnuta JM, Schaffer JL, Callaghan JJ, Mont MA, Kamath AF, Krebs VE, Ramkumar PN. Artificial intelligence and machine learning in lower extremity arthroplasty: a review. J Arthroplasty. 2019;34(10):2201-3.

19. Devana SK, Shah AA, Lee C, Roney AR, van der Schaar M, SooHoo NF. A novel, potentially universal machine learning algorithm to predict complications in total knee arthroplasty. Arthroplast Today. 2021;10: 135-43.

20. Jo C, Ko S, Shin WC, Han HS, Lee MC, Ko T, Ro DH. Transfusion after total knee arthroplasty can be predicted using the machine learning algorithm. Knee Surg Sports Traumatol Arthrosc. 2020;28(6):1757-64.

21. Karhade AV, Schwab JH, Bedair HS. Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. J Arthroplasty. 2019;34(10):2272-7.e1.

22. Navarro SM, Wang EY, Haeberle HS, Mont MA, Krebs VE, Patterson BM, Ramkumar PN. Machine learning and primary total knee arthroplasty: patient forecasting for a patient-specific payment model. J Arthroplasty. 2018; 33(12):3617-23.

23. Ramkumar PN, Navarro SM, Haeberle HS, Karnuta JM, Mont MA, Iannotti JP, Patterson BM, Krebs VE. Development and validation of a machine learning algorithm after primary total hip arthroplasty: applications to length of stay and payment models. J Arthroplasty. 2019; 34(4):632-7.

24. Roche C, Kumar V, Overman S, Simovitch R, Flurin PH, Wright T, Routman H, Teredesai A, Zuckerman J. Validation of a machine learning–derived clinical metric to quantify outcomes after total shoulder arthroplasty. J Shoulder Elbow Surg. 2021;30(10):2211-24.

25. Brochin RL, Zastrow RK, Patel AV, Parsons BO, Galatz LM, Flatow EL, Cagle PJ. Long-term clinical and radiographic outcomes of total shoulder arthroplasty in patients under age 60 years. J Shoulder Elbow Surg. 2022;31(6): S63-70.

26. Mattei L, Mortera S, Arrigoni C, Castoldi F. Anatomic shoulder arthroplasty: an update on indications, technique, results and complication rates. Joints. 2015;03(02):72-7.

27. Kumar V, Allen C, Overman S, Teredesai A, Simovitch R, Flurin PH, Wright T, Zuckerman J, Routman H, Roche C. Development of a predictive model for a machine learning–derived shoulder arthroplasty clinical outcome score. Semin Arthroplasty. 2022;32(2):226-37.

28. Polce EM, Kunze KN, Fu MC, Garrigues GE, Forsythe B, Nicholson GP, Cole BJ, Verma NN. Development of supervised machine learning algorithms for prediction of satisfaction at 2 years following total shoulder arthroplasty. J Shoulder Elbow Surg. 2021;30(6):e290-9.

29. Kumar V, Roche C, Overman S, Simovitch R, Flurin PH, Wright T, Zuckerman J, Routman H, Teredesai A. Using machine learning to predict clinical outcomes after shoulder arthroplasty with a minimal feature set. J Shoulder Elbow Surg. 2021;30(5):e225-36.

30. Gowd AK, Agarwalla A, Amin NH, Romeo AA, Nicholson GP, Verma NN, Liu JN. Construct validation of machine learning in the prediction of short-term postoperative complications following total shoulder arthroplasty. J Shoulder Elbow Surg. 2019;28(12):e410-21.

31. Simmons CS, Roche C, Schoch BS, Parsons M, Aibinder WR. Surgeon confidence in planning total shoulder arthroplasty improves after consulting a clinical decision support tool. Eur J Orthop Surg Traumatol. 2022. doi:10.1007/s00590-022-03446-1

32. Matsen FA III, Antoniou J, Rozencwaig R, Campbell B, Smith KL. Correlates with comfort and function after total shoulder arthroplasty for degenerative joint disease. J Shoulder Elbow Surg. 2000;9(6):465-9.

33. Orfaly RM, Rockwood CA Jr, Esenyel CZ, Wirth MA. A prospective functional outcome study of shoulder arthroplasty for osteoarthritis with an intact rotator cuff. J Shoulder Elbow Surg. 2003;12(3):214-21.

34. Alsubheen SA, MacDermid JC, John Faber K, James Overend T. Factors predicting postoperative range of motion and muscle strength one year after shoulder arthroplasty. Arch Bone Jt Surg. 2021;9(4):399-405.

35. Kumar V, Schoch BS, Allen C, Overman S, Teredesai A, Aibinder W, Parsons M, Watling J, Ko JK, Gobbato B, Throckmorton T, Routman H, Roche C. Using machine learning to predict internal rotation after anatomic and reverse total shoulder arthroplasty. J Shoulder Elbow Surg. 2022;31(5):e234-45.

36. Kumar V, Roche C, Overman S, Simovitch R, Flurin PH, Wright T, Zuckerman J, Routman H, Teredesai A. What is the accuracy of three different machine learning techniques to predict clinical outcomes after shoulder arthroplasty? Clin Orthop Relat Res. 2020; 478(10):2351-63.

37. Gowd AK, Charles MD, Liu JN, Lalehzarian SP, Cabarcas BC, Manderle BJ, Nicholson GP, Romeo AA, Verma NN. Single Assessment Numeric Evaluation (SANE) is a reliable metric to measure clinically significant improvements following shoulder arthroplasty. J Shoulder Elbow Surg. 2019;28(11):2238-46.

38. Jacobs CA, Morris BJ, Sciascia AD, Edwards TB. Comparison of satisfied and dissatisfied patients 2 to 5 years after anatomic total shoulder arthroplasty. J Shoulder Elbow Surg. 2016;25(7):1128-32.

39. Baettig SJ, Wieser K, Gerber C. Determinants of patient satisfaction following reconstructive shoulder surgery. BMC Musculoskelet Disord. 2017;18(1):458.

40. Berglund DD, Damodar D, Vakharia RM, Moeller EA, Giveans MR, Horn B, Mijic D, Levy JC. Predicting outstanding results after anatomic total shoulder arthroplasty using percentage of maximal outcome improvement. J Shoulder Elbow Surg. 2019; 28(2):349-56.

41. DeVito P, Damodar D, Berglund D, Vakharia R, Moeller EA, Giveans MR, Horn B, Malarkey A, Levy JC. Predicting outstanding results after reverse shoulder arthroplasty using percentage of maximal outcome improvement. J Shoulder Elbow Surg. 2019;28(7):1223-31.

42. Rauck RC, Ruzbarsky JJ, Swarup I, Gruskay J, Dines JS, Warren RF, Dines DM, Gulotta LV. Predictors of patient satisfaction after reverse shoulder arthroplasty. J Shoulder Elbow Surg. 2020;29(3):e67-74.

43. Angst F, Schwyzer H-K, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and Western Ontario Shoulder Instability Index (WOSI). Arthritis Care Res. 2011;63(suppl 11): S174-88.

44. McLendon PB, Christmas KN, Simon P, Plummer OR, Hunt A, Ahmed AS, Mighell MA, Frankle MA. Machine learning can predict level of improvement in shoulder arthroplasty. JB JS Open Access. 2021;6(1):e20.00128.

45. Tashjian RZ, Hung M, Keener JD, Bowen RC, McAllister J, Chen W, Ebersole G, Granger EK, Chamberlain AM. Determining the minimal clinically important difference for the American Shoulder and Elbow Surgeons score, Simple Shoulder Test, and visual analog scale (VAS) measuring pain after shoulder arthroplasty. J Shoulder Elbow Surg. 2017;26(1):144-8.

46. Lopez CD, Constant M, Anderson MJJ, Confino JE, Heffernan JT, Jobin CM. Using machine learning methods to predict nonhome discharge after elective total shoulder arthroplasty. JSES Int. 2021;5(4):692-8.

47. Devana SK, Shah AA, Lee C, Jensen AR, Cheung E, van der Schaar M, SooHoo NF. Development of a machine learning algorithm for prediction of complications and unplanned readmission following primary anatomic total shoulder replacements. J Shoulder Elb Arthroplast. 2022;6:247154922210754.

48. Devana SK, Shah AA, Lee C, Gudapati V, Jensen AR, Cheung E, Solorzano C, van der Schaar M, SooHoo NF. Development of a machine learning algorithm for prediction of complications and unplanned readmission following reverse total shoulder arthroplasty. J Shoulder Elb Arthroplast. 2021;5: 24715492211038172.

49. Arvind V, London DA, Cirino C, Keswani A, Cagle PJ. Comparison of machine learning techniques to predict unplanned readmission following total shoulder arthroplasty. J Shoulder Elbow Surg. 2021;30(2):e50-9.

50. Gowd AK, Agarwalla A, Beck EC, Rosas S, Waterman BR, Romeo AA, Liu JN. Prediction of total healthcare cost following total shoulder arthroplasty utilizing machine learning. J Shoulder Elbow Surg. 2022;31(12): 2449-56.

51. Karnuta JM, Churchill JL, Haeberle HS, Nwachukwu BU, Taylor SA, Ricchetti ET, Ramkumar PN. The value of artificial neural networks for predicting length of stay, discharge disposition, and inpatient costs after anatomic and reverse shoulder arthroplasty. J Shoulder Elbow Surg. 2020; 29(11):2385-94.

52. Andrews LB, Stocking C, Krizek T, Gottlieb L, Krizek C, Vargish T, Siegler M. An alternative strategy for studying adverse events in medical care. Lancet. 1997;349(9048):309-13.

53. Sivari E, Güzel MS, Bostanci E, Mishra A. A novel hybrid machine learning based system to classify shoulder implant manufacturers. Healthcare. 2022;10(3):580.

54. Yi PH, Kim TK, Wei J, Li X, Hager GD, Sair HI, Fritz J. Automated detection and classification of shoulder arthroplasty models using deep learning. Skeletal Radiol. 2020;49(10):1623-32.

55. Beyaz S, Açıcı K, Sümer E. Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches. Jt Dis Relat Surg. 2020;31(2):175-83.

56. Han Y, Wang G. Skeletal bone age prediction based on a deep residual network with spatial transformer. Comput Methods Programs Biomed. 2020;197:105754.

57. Üreten K, Arslan T, Gültekin KE, Demir AND, Özer HF, Bilgili Y. Detection of hip osteoarthritis by using plain pelvic radiographs with deep learning methods. Skeletal Radiol. 2020;49(9): 1369-74.

58. Sultan H, Owais M, Choi J, Mahmood T, Haider A, Ullah N, Park KR. Artificial intelligence-based solution in personalized computer-aided arthroscopy of shoulder prostheses. J Pers Med. 2022;12(1):109.

59. Geng EA, Cho BH, Valliani AA, Arvind V, Patel AV, Cho SK, Kim JS, Cagle PJ. Development of a machine learning algorithm to identify total and reverse shoulder arthroplasty implants from X-ray images. J Orthop. 2023;35:74-8.

60. Reid JJ, Kunkle BF, Greene AT, Eichinger JK, Friedman RJ. Variability and reliability of 2-dimensional vs. 3-dimensional glenoid version measurements with 3-dimensional preoperative planning software. J Shoulder Elbow Surg. 2022;31(2):302-9.

61. Lilley BM, Lachance A, Peebles AM, Powell SN, Romeo AA, Denard PJ, Provencher CMT. What is the deviation in 3D preoperative planning software? A systematic review of concordance between plan and actual implant in reverse total shoulder arthroplasty. J Shoulder Elbow Surg. 2022;31(5):1073-82.

62. Werner BS, Hudek R, Burkhart KJ, Gohlke F. The influence of three-dimensional planning on decision-making in total shoulder arthroplasty. J Shoulder Elbow Surg. 2017;26(8): 1477-83.