JSES International 5 (2021) 692–698

Contents lists available at ScienceDirect

JSES International

journal homepage: www.jsesinternational.org

Using machine learning methods to predict nonhome discharge after elective total shoulder arthroplasty

Cesar D. Lopez, BS*, Michael Constant, BA, Matthew J.J. Anderson, MD, Jamie E. Confino, MD, John T. Heffernan, MD, Charles M. Jobin, MD

New York-Presbyterian/Columbia University Irving Medical Center, New York, NY, USA

ARTICLE INFO

Keywords:
ACS-NSQIP
Machine learning
Artificial intelligence
Artificial neural network
Deep learning
Total shoulder arthroplasty
Nonhome discharge
Discharge disposition

Level of evidence: Level III, Retrospective Case-Control Design Using Large Database, Prognosis Study

Background: Machine learning has shown potential in accurately predicting outcomes after orthopedic surgery, thereby allowing for improved patient selection, risk stratification, and preoperative planning. This study sought to develop machine learning models to predict nonhome discharge after total shoulder arthroplasty (TSA).

Methods: The American College of Surgeons National Surgical Quality Improvement Program database was queried for patients who underwent elective TSA from 2012 to 2018. Boosted decision tree and artificial neural networks (ANN) machine learning models were developed to predict non-home discharge and 30-day postoperative complications. Model performance was measured using the area under the receiver operating characteristic curve (AUC) and overall accuracy (%). Multivariate binary logistic regression analyses were used to identify variables that were significantly associated with the predicted outcomes.

Results: There were 21,544 elective TSA cases identified in the National Surgical Quality Improvement Program registry from 2012 to 2018 that met inclusion criteria. Multivariate logistic regression identified several variables associated with increased risk of nonhome discharge including female sex (odds ratio [OR] = 2.83; 95% confidence interval [CI] = 2.53-3.17; $P < .001$), age older than 70 years (OR = 3.19; 95% CI = 2.86-3.57; $P < .001$), American Society of Anesthesiologists classification 3 or greater (OR = 2.70; 95% CI = 2.41-2.03; $P < .001$), prolonged operative time (OR = 1.38; 95% CI = 1.20-1.58; $P < .001$), as well as history of diabetes (OR = 1.56; 95% CI = 1.38-1.75; $P < .001$), chronic obstructive pulmonary disease (OR = 1.71; 95% CI = 1.46-2.01; $P < .001$), congestive heart failure (OR = 2.65; 95% CI = 1.72-4.01; $P < .001$), hypertension (OR = 1.35; 95% CI = 1.20-1.52; $P = .004$), dialysis (OR = 3.58; 95% CI = 2.01-6.39; $P = .002$), wound infection (OR = 5.67; 95% CI = 3.46-9.29; $P < .001$), steroid use (OR = 1.43; 95% CI = 1.18-1.74; $P = .010$), and bleeding disorder (OR = 1.84; 95% CI = 1.45-2.34; $P < .001$). The boosted decision tree model for predicting nonhome discharge had an AUC of 0.788 and an overall accuracy of 90.3%. The ANN model for predicting nonhome discharge had an AUC of 0.851 and an overall accuracy of 89.9%. For predicting the occurrence of 1 or more postoperative complications, the boosted decision tree model had an AUC of 0.795 and an overall accuracy of 95.5%. The ANN model yielded an AUC of 0.788 and an overall accuracy of 92.5%.

Conclusions: Both the boosted decision tree and ANN models performed well in predicting nonhome discharge with similar overall accuracy, but the ANN had higher discriminative ability. Based on the findings of this study, machine learning has the potential to accurately predict nonhome discharge after elective TSA. Surgeons can use such tools to guide patient expectations and to improve preoperative discharge planning, with the ultimate goal of decreasing hospital length of stay and improving cost-efficiency.

The demand for total shoulder arthroplasty (TSA) continues to rise throughout the United States, driven in large part by an aging population and expanding indications for reverse TSA.[12] Concurrently, efforts to curtail rising healthcare expenditures have intensified, with an emphasis on reducing inpatient costs associated with surgical procedures.[14,20,21,25,28-30] Recent advancements

No institutional review board approval was required for this retrospective study.

* Corresponding author: Cesar D. Lopez, BS, 622 W. 168th St. PH-11, New York, NY 10032, USA.

E-mail address: cdl2143@columbia.edu (C.D. Lopez).

https://doi.org/10.1016/j.jseint.2021.02.011

in regional anesthesia, minimally invasive surgical techniques, early mobilization with physical therapy, and multimodal analgesia have successfully reduced average length of stay (LOS) without compromising patient outcomes for many procedures,[7-9,15,17] but the recent shift toward bundled payment models further incentives surgeons and hospitals to reduce costs associated with postoperative care while limiting complications, readmissions, and reoperations.

Although the majority of patients with TSA are discharged home postoperatively, some patients require continued medical care or additional rehabilitation at postacute care facilities (eg, skilled nursing facilities, inpatient rehabilitation centers). Discharge to such facilities is frequently delayed leading to prolonged hospital LOS, which results in greater costs and has been associated with increased patient morbidity and mortality.[5,11,24,37] In a study of more than 1000 prospectively followed up hospital admissions, Andrews et al[2] found that patients with longer hospital stays had more adverse events on average after controlling for comorbidities; the likelihood of adverse events increased by nearly 6% for each additional day in the hospital. Accordingly, expediting the discharge process to postacute facilities has the potential to limit costs while decreasing patient morbidity and mortality. One tactic is to identify patients likely to require nonhome discharge, so that the necessary arrangements can be made preoperatively.

Machine learning (ML) is a form of artificial intelligence in which algorithms and statistical models automatically learn and improve by identifying patterns and complex relationships in large data sets, with the ultimate goal of making decisions using minimal human intervention.[3,13] Within orthopedic surgery, artificial intelligence /ML has proven beneficial in surgical risk stratification and preoperative optimization, outcome prediction, diagnostics, cost-efficiency analysis, and risk-adjusted insurance reimbursement models.[10,27,36] Among patients who underwent spine surgery, ML has successfully been used to accurately predict nonhome discharge.[19,33,34] Sivasundaram et al recently developed a statistical nomogram based on preoperative patient characteristics to predict discharge disposition after TSA, but no study has used ML for this purpose. As such, the aim of this study was to develop ML models based on data from a large national surgical registry to predict nonhome discharge after TSA.

Materials and methods

Data source

This retrospective study used population-level data from the American College of Surgeons National Surgical Quality Improvement Program (NSQIP) registry. The NSQIP data set contains deidentified information, including patient demographics, medical comorbidities, perioperative data, and 30-day postoperative outcomes for surgical patients in both inpatient and outpatient settings at more than 700 medical centers across the United States.[1] All patients who underwent TSA between 2012 and 2018 were identified using Current Procedural Terminology code 23472. These patients were then filtered to only include elective surgeries. Patients who underwent concomitant procedures (additional Current Procedural Terminology codes) or met criteria for sepsis/shock/systemic inflammatory response syndrome in the 48 hours before surgery were excluded. We included outpatient TSAs as part of a complete data set of elective TSAs, which may include outpatient TSA patients requiring nonhome discharge, especially if caused by unforeseen circumstances during or after surgery.

Preoperative patient demographic data including age, sex, race, body mass index (body mass index [BMI], calculated from the recorded height and weight), American Society of Anesthesiologists (ASA) classification, history of smoking, diabetes, chronic obstructive pulmonary disease (COPD), congestive heart failure, dyspnea at rest or with moderate exertion, steroid use, hypertension requiring medication, renal failure, dialysis requirement, disseminated cancer, use of steroids for a chronic condition, bleeding disorder, and need for a preoperative transfusion within 72 hours of surgery were abstracted. Perioperative data included anesthesia type (general, spinal, IV sedation, regional, other), surgery setting (inpatient vs. outpatient), and operative time (prolonged operative time defined as >120 minutes). The postoperative outcomes of interest were LOS, discharge destination, and occurrence of any complication within the first 30 days after surgery. Discharge destination was simplified to home discharge vs. nonhome discharge (eg, rehabilitation facility, skilled nursing facility, hospice, acute care facility). Postoperative complications assessed included surgical site infection, anemia requiring transfusion, deep vein thrombosis/pulmonary embolism, urinary tract infection, acute renal failure, sepsis, intubation-related complication, pneumonia, myocardial infarction, cerebrovascular event, cardiac arrest, and unplanned return to the operating room. Patients with missing data were excluded from analysis.

Statistical analysis

Descriptive statistics were calculated for all continuous variables and comparisons were made using unpaired Student's t-tests and one-way analysis of variance testing. A binary logistic regression model was used to calculate odds ratios (ORs) of nonhome discharge based on patient characteristics including age ($\geq$70 vs. < 70 years), race (white vs. non-white), BMI ($\geq$30 vs.<30), ASA classification (>2 vs. $\leq$2), anesthesia type (regional vs. general), diabetes, smoking history and other baseline patient characteristics. Results of the logistic regressions were reported as OR with 95% confidence intervals (CIs). Patients with missing data were excluded from the study. All statistical analysis was conducted on Stata, version 16.1 (Stata Corp., *College Station, TX, USA*). Statistical significance was defined as $P < .05$. Institutional review board approval was not required for this study.

Development of ML models

The data abstracted from the NSQIP registry were used to develop 2 types of ML models: a boosted decision tree model and an artificial neural networks (ANN) model. All model development and analysis were performed using the TensorFlow Python open-source coding platform (Google Brain, Alphabet Inc., *Mountain View, CA, USA*). Patients were randomly divided into a training set cohort (80%) and a testing set cohort (20%). We ensured that the testing data set (20%) was large enough to yield statistically meaningful results and was representative of the data set as a whole.

With less training data, there is greater variance of the model's parameter estimates, and with less testing data, the model's performance statistic will have greater variance. Therefore, the 80/20 data split ensures that both variance values are as low as possible, and many computer scientists refer to the Pareto principle for the 80/20 split. The overall data were large enough to ensure that the model did not match the training data so closely that it fails to make predictions on new data (testing set), which is also known as overfitting.

Patient data in the training set were used to develop and refine the ML models. For each incorrect prediction, the model self-calibrated through a process of reiterative algorithm refinement until optimal accuracy was achieved. Patient data in the testing set were used to evaluate model accuracy and performance.

Table I
Summary of patient demographics and medical comorbidities.

Predictive factors	Home discharge	Nonhome discharge	P value	All TSA
Women (%)	52.9%	77.4%	<.001	55.3%
Average Age (yr)	68.4	75.4	<.001	69.1
BMI	31.1	31.5	.005	31.1
Diabetes (%)	16.7%	24.6%	<.001	17.5%
Smoke (%)	11.1%	8.0%	<.001	10.8%
Dyspnea (%)	5.9%	13.5%	<.001	6.6%
COPD (%)	6.1%	12.1%	<.001	6.7%
CHF (%)	0.4%	1.9%	<.001	0.5%
Hypertension (%)	65.6%	78.0%	<.001	66.8%
Renal failure (%)	0.0%	0.0%	.353	0.0%
Dialysis (%)	0.3%	1.0%	<.001	0.4%
Cancer (%)	0.21%	0.33%	.231	0.2%
Wound infection (%)	0.2%	1.7%	<.001	0.4%
Steroid use (%)	4.6%	7.3%	<.001	4.9%
Weight loss (%)	0.17%	0.33%	.097	0.2%
Bleeding disorder (%)	2.3%	5.0%	<.001	2.5%
Transfusion (%)	0.1%	0.3%	.012	0.1%
Wound class > 1 (%)	0.8%	0.5%	.163	0.8%
ASA class > 2 (%)	53.1%	78.2%	<.001	55.6%

ASA, American Society of Anesthesiologists; *BMI*, body mass index; *CHF*, congestive heart failure; *COPD*, chronic obstructive pulmonary disease; *TSA*, total shoulder arthroplasty.

Categorical nonbinary variables (eg, race, ASA classification, anesthesia type) were incorporated into the models using 1-hot encoding. Continuous variables, such as age and BMI, were converted into logarithmic variables to normalize the data and minimize bias. Imbalanced data were managed using oversampling of the underrepresented variable.

The sensitivity and specificity were calculated for each model and used to develop a receiver operating characteristic curve. The area under the receiver operating characteristic curve (AUC) was then calculated as a measure of the model's discriminative ability, as has been carried out previously.[6,32] In this context, AUC values can range from 0.50 to 1, with a greater AUC signifying greater predictive capacity. A model with an AUC of 1.0 is a perfect discriminator, 0.90 to 0.99 is considered excellent, 0.80 to 0.89 is good, 0.70 to 0.79 is fair, and 0.51 to 0.69 is poor.[22] Overall model accuracy (%) was calculated by adding the number of true positives (correct predictions of outcome occurrence) and true negatives (correct prediction of outcome nonoccurrence) and dividing by the total sample size.

Results

The NSQIP registry contained 21,544 patients who underwent elective TSA between 2012 and 2018 and met all inclusion criteria (Table I). The mean age was 69.1 years (standard deviation 9.5 years), and 55.3% of patients were women. The majority of patients underwent TSA in the inpatient setting (92.3%). In terms of the primary form of anesthesia, 96.8% of patients received general anesthesia, 1.7% received regional anesthesia, and 0.9% received IV sedation (monitored anesthesia care). Operative time averaged 109.9 minutes (standard deviation 44.4 minutes), and mean hospital LOS was 1.7 days (standard deviation 2.2 days). Regarding discharge destination, 9.7% of patients underwent nonhome discharge. Overall, 4.5% of patients experienced 1 or more complications in the 30 days immediately after the surgery. Readmission and reoperation data were available for 16,757 patients (77.8% of total patients). Within this subset of patients, the 30-day readmission rate was 3.6%, with 2.5% of patients being readmitted for a TSA-related cause (defined using Current Procedural Terminology codes related to index procedure) and 1.3% of patients required reoperation related to TSA (Table II).

Direct comparison of patients who were discharged home vs. nonhome discharge patients revealed significant differences in terms of mean age (68.4 vs. 75.4 years, respectively; $P < .001$), sex (52.9% vs. 77.4% women, respectively; $P < .001$), average BMI (31.1 vs. 31.5, respectively; $P = .005$), and surgery setting (91.7% vs. 97.7% inpatient, respectively; $P < .001$) (Table I). Mean operative time was 109.9 vs. 110.0 minutes ($P = .925$), mean LOS was 1.6 vs. 3.5 days ($P < .001$), and the 30-day complication rate was 3.6% vs. 12.1% ($P < .001$) for home discharge and nonhome discharge patients, respectively (Table II). The overall readmission rate was 3.2% vs. 6.8% ($P < .001$), the TSA-related readmission rate was 2.3% vs. 4.5% ($P < .001$), and the reoperation rate was 1.2% vs. 1.8% ($P = .024$) for home discharge and nonhome discharge patients, respectively (Table II).

Multivariate binary logistic regression revealed that several variables were associated with greater odds of nonhome discharge including female sex (OR = 2.83; 95% CI = 2.53 to 3.17; $P < .001$), age older than 70 years (OR = 3.19; 95% CI = 2.86 to 3.57; $P < .001$), ASA classification 3 and greater (OR = 2.70; 95% CI = 2.41 to 3.03; $P < .001$), inpatient TSA (OR = 3.50; 95% CI = 2.58 to 4.73; $P < .001$), prolonged operative time (OR = 1.38; 95% CI = 1.20 to 1.58; $P < .001$), as well as history of diabetes (OR = 1.56; 95% CI = 1.38 to 1.75; $P < .001$), COPD (OR = 1.71; 95% CI = 1.46 to 2.01; $P < .001$), congestive heart failure (OR = 2.65; 95% CI = 1.72 to 4.01; $P < .001$), hypertension (OR = 1.35; 95% CI = 1.20 to 1.52; $P = .004$), dialysis (OR = 3.58; 95% CI = 2.01 to 6.39; $P = .002$), wound infection (OR = 5.67; 95% CI = 3.46 to 9.29; $P < .001$), steroid use (OR = 1.43; 95% CI = 1.18 to 1.74; $P = .010$), and bleeding disorder (OR = 1.84; 95% CI = 1.45 to 2.34; $P < .001$) (Table III). Similarly, variables associated with increased odds of experiencing a postoperative complication were female sex (OR = 1.38; 95% CI = 1.20 to 1.59; $P < .001$), age older than 70 years (OR = 1.61; 95% CI = 1.39 to 1.86; $P < .001$), ASA classification 3 and greater (OR = 2.40; 95% CI = 2.04 to 2.81; $P < .001$), prolonged operative time (OR = 2.36; 95% CI = 2.01 to 2.77; $P < .001$), as well as history of diabetes (OR = 1.45; 95% CI = 1.22 to 1.71; $P < .001$), smoking (OR = 1.25; 95% CI = 1.01 to 1.56; $P = .042$), COPD (OR = 1.86; 95% CI = 1.50 to 2.31; $P < .001$), congestive heart failure (OR = 2.39; 95% CI = 1.38 to 4.12; $P = .002$), dialysis (OR = 3.01; 95% CI = 1.51 to 6.00; $P = .002$), cancer (OR = 4.03; 95% CI = 1.82 to 8.93; $P = .001$), and bleeding disorder (OR = 2.42; 95% CI = 1.80 to 3.24; $P < .001$) (Table IV).

Table II
Summary of perioperative and postoperative outcomes.

Outcomes	Home discharge	Nonhome discharge	*P* value	All TSA
Inpatient TSA (%)	91.7%	97.7%	<.001	92.3%
Avg. operative time (min)	109.9	110.0	.925	109.9
Average LOS (d)	1.6	3.5	<.001	1.8
Any complication (%)	3.6%	12.1%	<.001	4.5%
Reoperations (%)	1.2%	1.8%	.024	1.3%
Readmissions (%)	3.2%	6.8%	<.001	3.6%
TSA-related readmissions (%)	2.3%	4.5%	<.001	2.5%

LOS, length of stay; *TSA*, total shoulder arthroplasty.

Table III
Factors associated with greater odds of non-home discharge, on multivariate logistic regression analysis.

Variable	Odds ratio	*P* value	[95% confidence interval]	
Preoperative factors				
Sex (female)	2.831	**<.001**	2.529	3.170
Race (white)	1.160	.166	0.940	1.431
Age > 70 yr	3.193	**<.001**	2.858	3.567
BMI > 30 (obese)	1.016	.756	0.919	1.123
Diabetes	1.556	**<.001**	1.383	1.751
Smoking	0.968	.728	0.804	1.165
COPD	1.709	**<.001**	1.455	2.006
CHF	2.653	**<.001**	1.718	4.097
HTN	1.348	**.004**	1.197	1.517
Dialysis	3.580	**.002**	2.007	6.385
Cancer	1.438	.434	0.579	3.569
Wound infection	5.669	**<.001**	3.459	9.290
Steroid use	1.429	**.010**	1.177	1.735
Bleeding disorder	1.839	**<.001**	1.449	2.335
ASA class > 2	2.703	**<.001**	2.412	3.029
Perioperative factors				
Inpatient TSA	3.496	**<.001**	2.584	4.729
Op. time > 150 min	1.378	**<.001**	1.200	1.583
Nongeneral anesthesia	1.125	.411	0.850	1.488

ASA, American Society of Anesthesiologists; *BMI*, body mass index; *CHF*, congestive heart failure; *COPD*, chronic obstructive pulmonary disease; *HTN*, hypertension; *TSA*, total shoulder arthroplasty.
Bolded: statistically significant values.

Table IV
Factors associated with greater odds of 30-day complication (any), on multivariate logistic regression analysis.

Variable	Odds ratio	*P* value	[95% Confidence interval]	
Preoperative factors				
Sex (female)	1.381	**<.001**	1.198	1.593
Race (white)	1.024	.869	0.777	1.349
Age > 70 yr	1.606	**<.001**	1.387	1.861
BMI > 30 (obese)	0.860	**.037**	0.746	0.991
Diabetes	1.448	**<.001**	1.224	1.712
Smoking	1.254	**.042**	1.008	1.560
COPD	1.863	**<.001**	1.504	2.306
CHF	2.387	**.002**	1.382	4.122
HTN	1.040	.629	0.888	1.218
Dialysis	3.006	**.002**	1.506	6.001
Cancer	4.030	**.001**	1.818	8.934
Wound infection	1.435	.386	0.634	3.248
Steroid use	1.220	.165	0.922	1.614
Bleeding disorder	2.416	**<.001**	1.804	3.235
ASA class > 2	2.397	**<.001**	2.044	2.811
Perioperative factors				
Inpatient TSA	1.333	.051	0.999	1.779
Op. time > 150 min	2.359	**<.001**	2.008	2.773
Non-general anesthesia	0.840	.443	0.539	1.311

ASA, American Society of Anesthesiologists; *BMI*, body mass index; *CHF*, congestive heart failure; *COPD*, chronic obstructive pulmonary disease; *HTN*, hypertension; *TSA*, total shoulder arthroplasty.
Bolded: statistically significant values.

To develop the ML models, the data were randomly divided into a training cohort of 17,235 patients and a testing cohort of 4309 patients. On testing the boosted decision tree model for predicting nonhome discharge, the AUC was 0.788 and the overall accuracy was 90.3% (Fig. 1*A*). The ANN model for predicting nonhome discharge had an AUC of 0.851 and an overall accuracy of 89.9% (Fig. 1*B*). In terms of predicting the occurrence of 1 or more postoperative complications, the boosted decision tree model had an AUC of 0.795 and an overall accuracy of 95.5% (Fig. 2*A*). The ANN model yielded an AUC of 0.788 and an overall accuracy of 92.5% (Fig. 2*B*).

Discussion

This is the first study to successfully develop ML models for predicting nonhome discharge after elective TSA. The predictive capacity, as measured by AUC on the receiver operating characteristic curve, was fair for the boosted decision tree model (AUC of 0.788) and good for the ANN model (0.851). Both models were similar in terms of overall accuracy (90.3% for the boosted decision tree model vs. 89.9% for the ANN model). With regard to predicting the occurrence of 30-day postoperative complications, both models performed in the fair range (AUC of 0.795 for the boosted decision tree model vs. 0.788 for the ANN model) with an overall accuracy of 95.5% for the boosted decision tree model and 92.5% for the ANN model.

Overall, 9.7% of patients were discharged to a postacute care facility after elective TSA, which is comparable with the nonhome discharge rate observed by Sivasundaram et al[40] (11.5%). Several patient factors were associated with increased risk of nonhome discharge including female sex, age older than 70 years, ASA classification 3 or greater, and history of diabetes, hypertension, wound infection, steroid use, or a bleeding disorder. Similar associations have been reported in previous studies of TSA and total knee arthroplasty, including a systematic review by Berman et al in which female sex, older age, obesity, and reverse TSA were associated with nonhome discharge.[39,40]

Hospital discharge after any surgical procedure is a complex process that depends on coordination between numerous parties including the patient, the surgeon, the social worker, the care coordinator, the physical therapist, the nurse, and the hospital administrator. Countless factors can delay discharge including disagreement between any of the involved parties regarding the appropriate level of postacute care needed, patient preference for a specific facility, facility bed availability, and insurance approval. Whatever the reason, a delay in discharge results in a prolonged hospital stay, which not only leads to increased costs but has been associated with greater patient morbidity and mortality.[5,11,24,37] Interestingly, Menendez et al[31] found that prolonged hospitalizations after TSA were most commonly attributed to issues with insurance approval and lack of social support, often ending in patients being discharged to skilled nursing facilities. As such, the

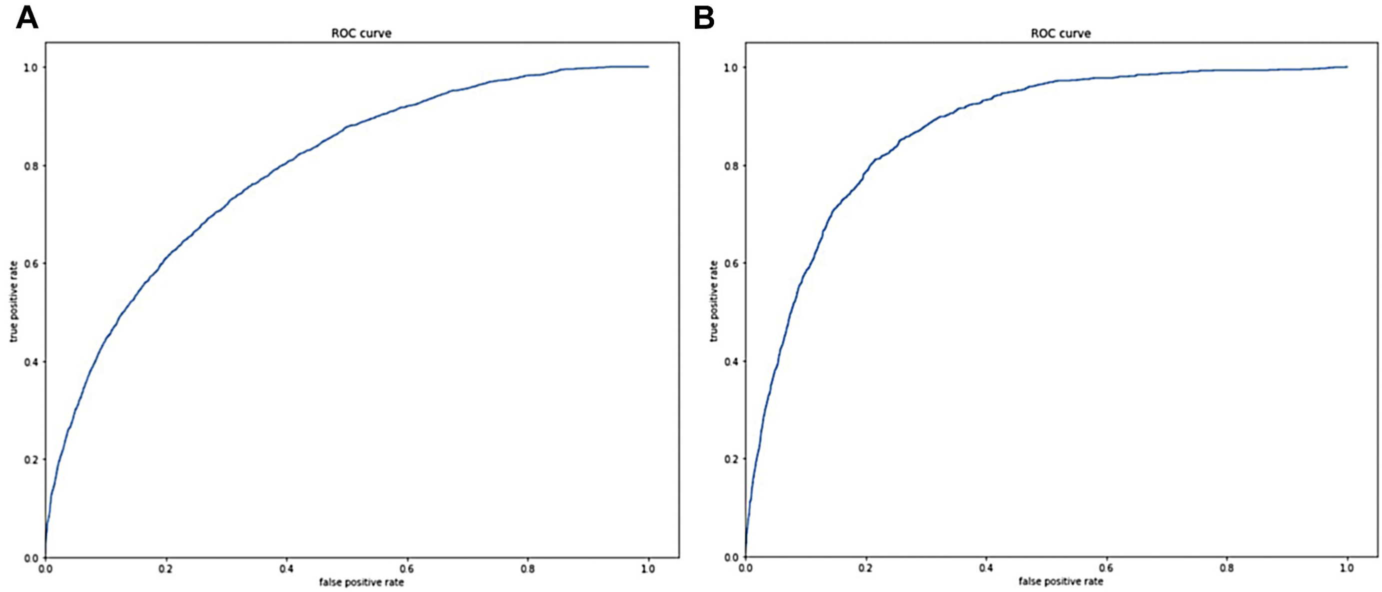

Figure 1 Area under the ROC curve of boosted decision tree (**A**) and artificial neural network (**B**) models of nonhome discharge. *ROC*, receiver operating characteristic.

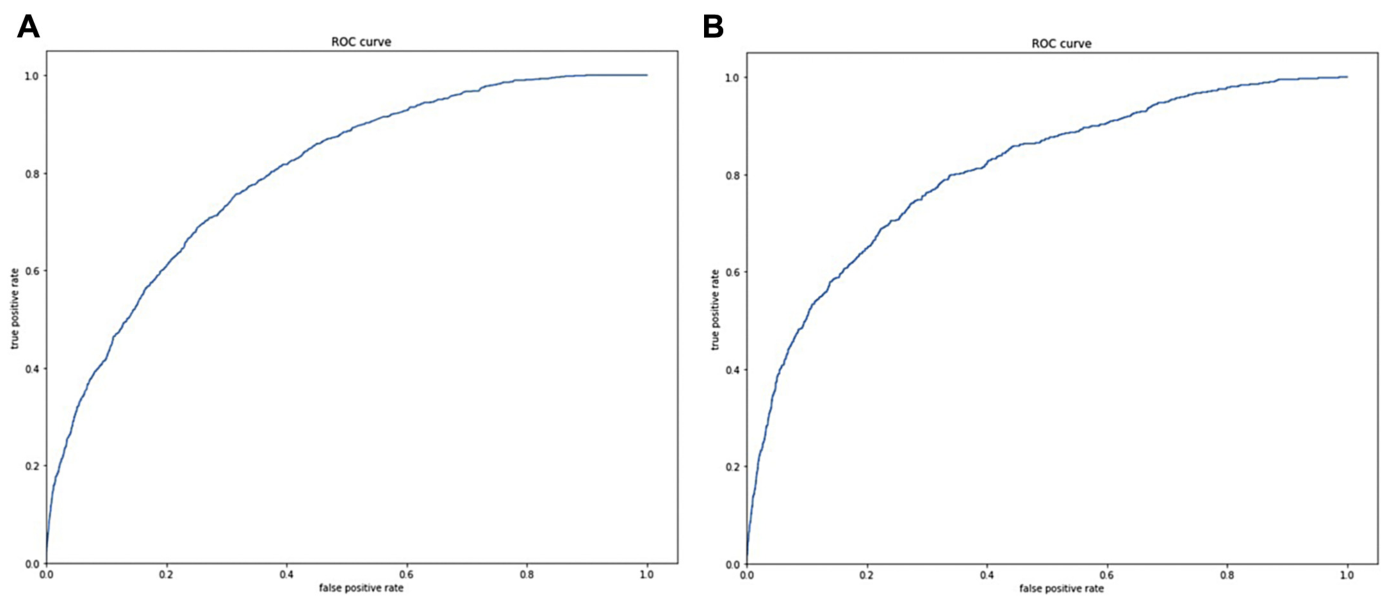

Figure 2 Area under the ROC curve of boosted decision tree (**A**) and artificial neural network (**B**) models of any 30-day complication. *ROC*, receiver operating characteristic.

ability to identify patients likely to require nonhome discharge after TSA affords surgeons the opportunity to make the necessary preparations beforehand, which can ultimately reduce hospital LOS, limit costs, and improve patient outcomes.

With the recent popularization of bundled payment models, healthcare facilities and providers are incentivized now more than ever to discharge patients in a safe and timely manner. As postoperative care is included under most bundled payment models, prolonged hospitalization directly diminishes reimbursement for facilities and providers. In this manner, models that can accurately identify patients at increased risk of nonhome discharge may allow for significant cost savings in that discharge planning can begin preoperatively, thereby reducing hospital LOS. For instance, Barsoum et al[4] demonstrated that use of a statistical model for predicting discharge to a postacute care facility after total joint arthroplasty resulted in a 0.9 day decrease in total hospital LOS.

It was somewhat expected that the ANN model had slightly better discriminative ability than the boosted decision tree model in predicting nonhome discharge. While both model types can be constructed based on variables with nonlinear relationships, neural networks consist of multiple "neural layers" that can process large amounts of data and share information using weighted connections that are optimized during the training process.[23,35] As a result, neural networks can achieve predictive performance, which provides an advantage in accurately modeling complex nonlinear relationships in high-volume data sets. Boosted decision tree models use a set of learned rules to make predictions but lack generalization to nonlinear interactions between variables or random events.[26] In this study, nonhome discharge was a more difficult outcome to predict because discharge destination can be influenced by a variety of variables not included in the NSQIP registry, such as ability to complete activities of daily living.

While ML has the potential to reduce healthcare costs and improve patient outcomes, serious ethical concerns have been raised. First, ML models could be used by insurance companies and providers to restrict access to care for patients at increased risk of costly postoperative complications and adverse events. Second, ML models developed from large national data sets may be inherently biased owing to historically inequitable healthcare data between racial and socioeconomic groups.[16,18] Given the well-documented systemic racial and socioeconomic disparities that exist in the US healthcare system,[16] administrators must be aware of these potential biases and generalizability when implementing ML technology in clinical practice. Third, the privacy risks associated with managing the large amounts of sensitive data necessary to build ML models remain a persistent challenge.

This study has several limitations that merit further discussion. Most notably, the ML tools developed in this study require external validation before definitive conclusions can be made regarding their efficacy and utility in the clinical setting. In addition, there are inherent limitations to using large registries, such as NSQIP, including coding errors, missing data, and inaccurate information. Future research should attempt to replicate this study's findings using other data sets, especially as the number of TSA procedures performed each year continues to rise. In addition, although ASA classification, which accounts for multiple patient factors, was included in our analysis, the effect of combined comorbidities (ie, age and COPD history) on outcomes should be investigated in future studies. Finally, the models developed in this study oversimplified the hospital discharge process, only considering only home vs nonhome. In reality, numerous types of postacute care facilities exist, with significant variability in terms of acceptance requirements, insurance approval, and cost.[38]

Conclusions

In the midst of a transition to high-value, cost-conscious health care, ML has the potential to improve outcomes and reduce costs by identifying patients likely to experience adverse outcomes following surgical procedures. This information can be used to enhance patient selection, risk stratification, counseling, optimization, and preoperative planning. The models developed in this study demonstrate an application of ML in which data from a national registry was used to predict nonhome discharge after elective TSA. Based on the predictions of these models, surgeons can better identify candidates for outpatient surgery and improve preoperative discharge planning for patients requiring inpatient surgery, thereby reducing hospital length of stay, associated costs, and patient morbidity and mortality.

Disclaimers:

Funding: No funding was disclosed by the author(s).
Conflicts of interest: Charles M. Jobin reports being a paid consultant, presenter, speaker and research support from Acumed, LLC., paid consulting fees from CFO LLC, paid consultant fees from Depuy, A Johnson & Johnson Company, paid consultant fees from Integral Life Sciences, paid presenter or speaker fees from Tornier, paid consultant fees from Wright Medical Technology and paid consultant, presenter, or speaker fees from Zimmer.

The other authors, their immediate family, and any research foundation with which they are affiliated did not receive any financial payments or other benefits from any commercial entity related to the subject of this article

References

1. American College of Surgeons National Surgical Quality Improvement Program. User Guide for the 2017 ACS NSQIP Participant Use Data File (PUF) 2017. American College of Surgeons: Washington, DC.
2. Andrews LB, Stocking C, Krizek T, Gottlieb L, Krizek C, Vargish T, et al. An alternative strategy for studying adverse events in medical care. Lancet 1997;349:309-13.
3. Awan SE, Sohel F, Sanfilippo FM, Bennamoun M, Dwivedi G. Machine learning in heart failure: ready for prime time. Curr Opin Cardiol 2018;33:190-5. https://doi.org/10.1097/HCO.0000000000000491.
4. Barsoum WK, Murray TG, Klika AK, Green K, Miniaci SL, Wells BJ, et al. Predicting patient discharge disposition after total joint arthroplasty in the United States. J Arthroplasty 2010;25:885-92. https://doi.org/10.1016/j.arth.2009.06.022.
5. Benson RT, Drew JC, Galland RB. A waiting list to go home: an analysis of delayed discharges from surgical beds. Ann R Coll Surg Engl 2006;88:650-2. https://doi.org/10.1308/003588406X149246.
6. Biron DR, Sinha I, Kleiner JE, Aluthge DP, Goodman AD, Sarkar IN, et al. A Novel Machine Learning Model Developed to Assist in Patient Selection for Outpatient Total Shoulder Arthroplasty. J Am Acad Orthop Surg 2020;28:e580-5. https://doi.org/10.5435/JAAOS-D-19-00395.
7. Brolin TJ, Mulligan RP, Azar FM, Throckmorton TW. Neer Award 2016: Outpatient total shoulder arthroplasty in an ambulatory surgery center is a safe alternative to inpatient total shoulder arthroplasty in a hospital: a matched cohort study. J Shoulder Elbow Surg 2017;26:204-8. https://doi.org/10.1016/j.jse.2016.07.011.
8. Brolin TJ, Throckmorton TW. Outpatient Shoulder Arthroplasty. Orthop Clin North Am 2018;49:73-9. https://doi.org/10.1016/j.ocl.2017.08.011.
9. Cancienne JM, Brockmeier SF, Gulotta LV, Dines DM, Werner BC. Ambulatory Total Shoulder Arthroplasty: A Comprehensive Analysis of Current Trends, Complications, Readmissions, and Costs. J Bone Joint Surg Am 2017;99:629-37. https://doi.org/10.2106/JBJS.16.00287.
10. Chung SW, Han SS, Lee JW, Oh KS, Kim NR, Yoon JP, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop 2018;89:468-73. https://doi.org/10.1080/17453674.2018.1453714.
11. Costa AP, Poss JW, Peirce T, Hirdes JP. Acute care inpatients with long-term delayed-discharge: evidence from a Canadian health region. BMC Health Serv Res 2012;12:172. https://doi.org/10.1186/1472-6963-12-172.
12. Day JS, Lau E, Ong KL, Williams GR, Ramsey ML, Kurtz SM. Prevalence and projections of total shoulder and elbow arthroplasty in the United States to 2015. J Shoulder Elbow Surg 2010;19:1115-20. https://doi.org/10.1016/j.jse.2010.02.009.
13. Deo RC. Machine Learning in Medicine. Circulation 2015;132:1920-30. https://doi.org/10.1161/CIRCULATIONAHA.115.001593.
14. Dietz N, Sharma M, Alhourani A, Ugiliweneza B, Wang D, Nuno MA, et al. Bundled Payment Models in Spine Surgery: Current Challenges and Opportunities, a Systematic Review. World Neurosurg 2019;123:177-83. https://doi.org/10.1016/j.wneu.2018.12.001.
15. Dunn JC, Lanzi J, Kusnezov N, Bader J, Waterman BR, Belmont PJ Jr. Predictors of length of stay after elective total shoulder arthroplasty in the United States. J Shoulder Elbow Surg 2015;24:754-9. https://doi.org/10.1016/j.jse.2014.11.042.
16. Feagin J, Bennefield Z. Systemic racism and U.S. health care. Soc Sci Med 2014;103:7-14. https://doi.org/10.1016/j.socscimed.2013.09.006.
17. Fournier MN, Brolin TJ, Azar FM, Stephens R, Throckmorton TW. Identifying appropriate candidates for ambulatory outpatient shoulder arthroplasty: validation of a patient selection algorithm. J Shoulder Elbow Surg 2019;28:65-70. https://doi.org/10.1016/j.jse.2018.06.017.
18. Gibney E. The battle for ethical AI at the world's biggest machine-learning conference. Nature 2020;577:609. https://doi.org/10.1038/d41586-020-00160-y.
19. Goyal A, Ngufor C, Kerezoudis P, McCutcheon B, Storlie C, Bydon M. Can machine learning algorithms accurately predict discharge to nonhome facility and early unplanned readmissions following spinal fusion? Analysis of a national surgical registry. J Neurosurg Spine 2019:1-11. https://doi.org/10.3171/2019.3.SPINE181367.
20. Grenda TR, Pradarelli JC, Thumma JR, Dimick JB. Variation in Hospital Episode Costs With Bariatric Surgery. JAMA Surg 2015;150:1109-15. https://doi.org/10.1001/jamasurg.2015.2394.
21. Guduguntla V, Syrjamaki JD, Ellimoottil C, Miller DC, Prager RL, Norton EC, et al. Drivers of Payment Variation in 90-Day Coronary Artery Bypass Grafting Episodes. JAMA Surg 2018;153:14-9. https://doi.org/10.1001/jamasurg.2017.2881.

22. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
23. Hazan H, Saunders DJ, Khan H, Patel D, Sanghavi DT, Siegelmann HT, et al. BindsNET: A Machine Learning-Oriented Spiking Neural Networks Library in Python. Front Neuroinform 2018;12:89. https://doi.org/10.3389/fninf.2018.00089.
24. Hwabejire JO, Kaafarani HM, Imam AM, Solis CV, Verge J, Sullivan NM, et al. Excessively long hospital stays after trauma are not related to the severity of illness: let's aim to the right target! JAMA Surg 2013;148:956-61. https://doi.org/10.1001/jamasurg.2013.2148.
25. Kahn EN, Ellimoottil C, Dupree JM, Park P, Ryan AM. Variation in payments for spine surgery episodes of care: implications for episode-based bundled payment. J Neurosurg Spine 2018;29:214-9. https://doi.org/10.3171/2017.12.SPINE17674.
26. Kostoglou K, Michmizos KP, Stathis P, Sakas D, Nikita KS, Mitsis GD. Classification and Prediction of Clinical Improvement in Deep Brain Stimulation From Intraoperative Microelectrode Recordings. IEEE Trans Biomed Eng 2017;64:1123-30. https://doi.org/10.1109/TBME.2016.2591827.
27. Lin CC, Ou YK, Chen SH, Liu YC, Lin J. Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. Injury 2010;41:869-73. https://doi.org/10.1016/j.injury.2010.04.023.
28. Malik AT, Deiparine S, Khan SN, Kim J, Yu E. Costs Associated With a 90-Day Episode of Care After Single-Level Anterior Lumbar Interbody Fusion. World Neurosurg 2020;135:e716-22. https://doi.org/10.1016/j.wneu.2019.12.117.
29. Malik AT, Xie J, Retchin SM, Phillips FM, Xu W, Yu E, et al. Primary single-level lumbar microdisectomy/decompression at a free-standing ambulatory surgical center vs. a hospital-owned outpatient department-an analysis of 90-day outcomes and costs. Spine J 2020;20:882-7. https://doi.org/10.1016/j.spinee.2020.01.015.
30. Malik AT, Yu E, Kim J, Khan SN. Understanding Costs in a 90-Day Episode of Care Following Posterior Spinal Fusions for Adolescent Idiopathic Scoliosis. World Neurosurg 2019;130:e535-41. https://doi.org/10.1016/j.wneu.2019.06.149.
31. Menendez ME, Baker DK, Fryberger CT, Ponce BA. Predictors of extended length of stay after elective shoulder arthroplasty. J Shoulder Elbow Surg 2015;24:1527-33. https://doi.org/10.1016/j.jse.2015.02.014.
32. Navarro SM, Wang EY, Haeberle HS, Mont MA, Krebs VE, Patterson BM, et al. Machine Learning and Primary Total Knee Arthroplasty: Patient Forecasting for a Patient-Specific Payment Model. J Arthroplasty 2018;33:3617-23. https://doi.org/10.1016/j.arth.2018.08.028.
33. Ogink PT, Karhade AV, Thio Q, Gormley WB, Oner FC, Verlaan JJ, et al. Predicting discharge placement after elective surgery for lumbar spinal stenosis using machine learning methods. Eur Spine J 2019;28:1433-40. https://doi.org/10.1007/s00586-019-05928-z.
34. Ogink PT, Karhade AV, Thio Q, Hershman SH, Cha TD, Bono CM, et al. Development of a machine learning algorithm predicting discharge placement after surgery for spondylolisthesis. Eur Spine J 2019;28:1775-82. https://doi.org/10.1007/s00586-019-05936-z.
35. Patel JL, Goyal RK. Applications of artificial neural networks in medical science. Curr Clin Pharmacol 2007;2:217-26. https://doi.org/10.2174/157488407781668811.
36. Ramkumar PN, Navarro SM, Haeberle HS, Karnuta JM, Mont MA, Iannotti JP, et al. Development and Validation of a Machine Learning Algorithm After Primary Total Hip Arthroplasty: Applications to Length of Stay and Payment Models. J Arthroplasty 2019;34:632-7. https://doi.org/10.1016/j.arth.2018.12.030.
37. Rosman M, Rachminov O, Segal O, Segal G. Prolonged patients' In-Hospital Waiting Period after discharge eligibility is associated with increased risk of infection, morbidity and mortality: a retrospective cohort analysis. BMC Health Serv Res 2015;15:246. https://doi.org/10.1186/s12913-015-0929-6.
38. Sigurdsson E, Siggeirsdottir K, Jonsson H Jr, Gudnason V, Matthiasson T, Jonsson BY. Early discharge and home intervention reduces unit costs after total hip replacement: results of a cost analysis in a randomized study. Int J Health Care Finance Econ 2008;8:181-92. https://doi.org/10.1007/s10754-008-9036-0.
39. Sivasundaram L, Heckmann N, Pannell WC, Alluri RK, Omid R, Hatch GF 3rd. Preoperative risk factors for discharge to a postacute care facility after shoulder arthroplasty. J Shoulder Elbow Surg 2016;25:201-6. https://doi.org/10.1016/j.jse.2015.07.028.
40. Sivasundaram L, Tanenbaum J, Trivedi N, Kim CY, Bafus BT, Hoyen HA, et al. Predicting Discharge Disposition After Shoulder Arthroplasty in the United States. J Am Acad Orthop Surg 2019;27:e444-50. https://doi.org/10.5435/JAAOS-D-17-00743.