



Using Predictive Modeling and Supervised Machine Learning to Identify Patients at Risk for Venous Thromboembolism Following Posterior Lumbar Fusion

Kevin Y. Wang, BS¹, Ijezie Ikwuezunma, BS, BA¹ ,
Varun Puvanesarajah, MD¹, Jacob Babu, MD¹, Adam Margalit, MD¹,
Micheal Raad, MD¹, and Amit Jain, MD¹ 

Abstract

Study Design: Retrospective review.

Objective: To use predictive modeling and machine learning to identify patients at risk for venous thromboembolism (VTE) following posterior lumbar fusion (PLF) for degenerative spinal pathology.

Methods: Patients undergoing single-level PLF in the inpatient setting were identified in the National Surgical Quality Improvement Program database. Our outcome measure of VTE included all patients who experienced a pulmonary embolism and/or deep venous thrombosis within 30-days of surgery. Two different methodologies were used to identify VTE risk: 1) a novel predictive model derived from multivariable logistic regression of significant risk factors, and 2) a tree-based extreme gradient boosting (XGBoost) algorithm using preoperative variables. The methods were compared against legacy risk-stratification measures: ASA and Charlson Comorbidity Index (CCI) using area-under-the-curve (AUC) statistic.

Results: 13,500 patients who underwent single-level PLF met the study criteria. Of these, 0.95% had a VTE within 30-days of surgery. The 5 clinical variables found to be significant in the multivariable predictive model were: age > 65, obesity grade II or above, coronary artery disease, functional status, and prolonged operative time. The predictive model exhibited an AUC of 0.716, which was significantly higher than the AUCs of ASA and CCI (all, $P < 0.001$), and comparable to that of the XGBoost algorithm ($P > 0.05$).

Conclusion: Predictive analytics and machine learning can be leveraged to aid in identification of patients at risk of VTE following PLF. Surgeons and perioperative teams may find these tools useful to augment clinical decision making risk stratification tool.

Keywords

spine, posterior lumbar fusion, venous thromboembolism, predictive modeling, machine learning

Introduction

Posterior lumbar fusion (PLF) is a commonly performed surgery for degenerative spinal diseases including lumbar spinal stenosis and spondylolisthesis¹. From 1998 to 2014, the utilization rate for spinal fusions has increased from 74 to 139 cases per 100,000 persons, with lumbar fusions being the primary driver of this trend.^{2–5} Despite increasing surgeon experience and more advanced techniques, complication rates have remained relatively stable over time.^{3,6} Among the commonly cited complications of PLF, venous thromboembolism (VTE) occurs in 0.13% to 3.7% of cases and is a notable source of

morbidity, mortality, and increased healthcare utilization.^{7–9} If not treated promptly, these thromboembolic complications can lead to cardiopulmonary compromise, myocardial infarction,

¹ Department of Orthopaedic Surgery, The Johns Hopkins Hospital, Baltimore, MD, USA

Corresponding Author:

Amit Jain, Department of Orthopaedic Surgery, The Johns Hopkins Hospital, 1800 Orleans St, Baltimore, MD 21287, USA.
Email: amitjain@jhmi.edu



cerebral ischemia, and even sudden death.¹⁰ Many risk factors have been previously identified, with a recent systematic review reporting elderly age, female gender, diabetes, chronic kidney disease, non-ambulatory functional status, prolonged operative time, and postoperative blood transfusion as the independent risk factors for VTE after spine surgery.⁹ While several studies have determined independent risk factors for VTE in spine surgery using conventional regression analysis, ultimately more comprehensive and implementable tools, such as predictive models and potentially machine learning algorithms, are needed in clinical practice.

Given the importance of improving patient outcomes and quality of care, as well as lowering costs associated with complications, there has been a focus recently on preoperative risk stratification in orthopedic surgery.^{11,12} A variety of risk stratification scales for other orthopedic surgeries have been proposed including for outpatient spine surgery, total joint arthroplasty, and open reduction and internal fixation of ankle fractures.¹³⁻¹⁵ At present, surgeons use a combination of clinical judgment and legacy risk-stratification metrics like American Society of Anesthesiologists (ASA) and Charlson Comorbidity Index (CCI) to identify patients who may be at risk of complications following spine surgery.¹⁶⁻¹⁸ However, to our knowledge, there are no standardized risk stratification tools for determining patients at risk of VTE following PLF.

The purpose of this study was to develop clinically useful tools using predictive analytics and machine learning to improve upon these existing methods for identifying patients at risk of postoperative VTE following PLF. We hypothesized that both the scoring tool and the machine learning model would be predictive of VTE, though the machine learning model would have a higher receiver operator area under the curve (AUC) metric due to its greater computational complexity relative to the scoring tool.

Methods

Data Source

The National Surgical Quality Improvement Program (NSQIP) database is a prospectively collected surgical registry containing data on more than 7-million patients from over 500 participating United States medical institutions. The registry includes information on patient demographics and comorbidities, surgical procedures, metrics of healthcare utilization (i.e. operative time, length of stay), and 30-day postoperative complications. The NSQIP database has been shown to have high validity and is commonly used for orthopedic research.¹⁹ After institutional review, our study was exempted from IRB review and informed patient consent.

Patient Population

Adult patients undergoing in-patient, single-level PLF from 2010 to 2017 were identified using Current Procedural Terminology (CPT) codes 22612, 22630, 22633. We specified

single-level PLF by excluding patients with CPT codes denoting add-on levels (22614, 22632, 22634). Patients with International Classification of Diseases Ninth Revision and Tenth Revision (ICD) diagnosis codes indicating revision surgery, trauma, or vertebral malignancy were excluded (N = 389). All codes used to define our cohort are displayed in Online Appendix 1. Patients with any missing data in the outcome measure or predictive variables were removed from the analysis (N = 1,174).

Outcome Measure and Predictive Variables

The primary outcome measure was VTE, which was defined as the presence of either deep vein thrombosis (DVT) or pulmonary embolism (PE) within 30 days of the index operation.

Variables tested in the initial multivariable logistic regression for inclusion into the predictive model and extreme gradient boosting (XGBoost) machine learning algorithm included: age, sex, body mass index (BMI), diabetes, congestive heart failure, coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), hypertension, renal failure, hematocrit, albumin, INR, bilirubin, current tobacco use, functional status, and operative time. Prolonged operative time was defined as above the 90th percentile. Per convention, obesity was categorized in terms of grades (grade I: BMI 31-35; grade II: BMI 36-40; grade III: ≥ 41).

Statistical Analysis—Predictive Model

A randomly split 80% of our dataset was used to train our models, and the remaining 20% was used to test our models. We performed multivariable logistic regression with backward elimination to determine the variables most predictive of VTE, and these variables were included in the predictive model. Based on nomogram analysis of our logistic regression, each variable was given a scaled weight based on its relative importance within the model. The nomogram was calibrated such that the maximum weight was 3. The final resulting model was then used to predict VTE in the test cohort, and the most effective cut-off value for the model was determined using Youden's index analysis. The methodologies used to construct our score-based model were designed and executed based on previous peer-reviewed methodologies.^{13-15,20}

Statistical Analysis—XGBoost

We again randomly split our dataset into 80% used to train our XGBoost models and 20% use to test our models. XGBoost is a form of supervised machine learning which uses decision trees to iteratively partition data from the training dataset into smaller, distinct subsets based on similar characteristics using the "dart" boosting method.^{21,22} We used a native feature importance algorithm built upon multivariate logistic regression to identify and include the top 10 most predictive variables for our outcome measure: BMI, operative time, hematocrit, age, albumin, CAD, gender, smoking status, COPD, and

functional status. We only included these 10 variables to minimize any bias from overfitting our model.²³ After the model was constructed using these 10 variables, predictions on new data in the test set are made based on the probability of a datapoint falling within a given subset having a particular outcome based on the outcomes of the other datapoints in that subset. We internally validated our XGBoost algorithm by applying it to the test dataset.

Statistical Analysis—Validation

AUC analysis was used to measure the effectiveness of the predictive model and the XGBoost algorithm. We performed similar AUC analysis on ASA and CCI legacy indices.^{24,25}

AUC scores were interpreted according to statistical convention: an AUC score of 0.5 indicates no discriminative ability, and 0.7-0.8 indicates good discrimination.²⁶ Significance was set at *P* = 0.05 for analyses. Our scoring system was developed and validated using STATA version 15.0 (Stata-Corp, College Station, TX). The XGBoost algorithm was constructed and validated using R, version 4.0.3 (R Project for Statistical Computing).

Results

Study Population

A total of 13,500 adult patients who underwent inpatient single-level PLF were identified in the NSQIP registry. The mean age was 50.6 years old, and 55.7% (7,516) were female (Table 1). Of the 13,500 patients in our final cohort, 0.95% (128) were found to have a VTE within 30-days of surgery. Detailed demographic and comorbidity data is shown in Table 1. Mean operative time was 193 minutes, and prolonged operative time was defined as >90th percentile (>304 minutes).

Predictive Model Parameters

For the 10,750 patients included in our training dataset, 5 variables were independently predictive of VTE on multivariable analysis: age over 65 (OR = 1.52, 95% CI [1.07, 2.17], *P* = 0.019), obesity grade II or higher (OR = 1.56, 95% CI [1.07, 2.26], *P* = 0.019), history of CAD (OR = 10.52, 95% CI [2.33, 47.59], *P* = 0.019), dependent functional status (OR = 2.40, 95% CI [1.16, 4.97], *P* < 0.001), and prolonged operative time (OR = 3.27, 95% CI [2.19, 4.87], *P* < 0.001). The univariate and multivariable regression analyses demonstrating the associations between VTE risk and each of these variables are displayed in Tables 2 and 3, respectively. These variables were used to create the predictive model, and nomogram analysis produced weighted point values for each variable. One point was assigned for: elderly, obesity, and dependent functional status. Two points were assigned for prolonged operative time, and 3 points were assigned for CAD.

Table 1. Patient Demographics and Comorbidities.

	N	%
Total	13,500	-
Age		
19-34	490	3.63%
35-49	2,146	15.90%
50-65	5,050	37.41%
65+	5,814	43.07%
BMI		
<30	7,244	53.66%
31-35	3,692	27.35%
36-40	1,537	11.39%
40+	1,027	7.61%
Sex		
Female	7,516	55.67%
Comorbidities		
Diabetes	2,797	20.72%
COPD	714	5.29%
CAD	19	0.14%
CHF	44	0.33%
Renal failure	7	0.05%
Functional status		
Independent	13,214	97.88%
Partially dependent	269	1.99%
Totally dependent	17	0.13%

Table 2. Univariate Analysis of Risk Factors for Venous Thromboembolism.

Category	Venous thromboembolism		
	N	%	<i>P</i> -value ^a
Total	128	-	-
Elderly (age > 65 years)	63	1.17%	0.032
Obesity (grade II or above)	44	1.33%	0.009
Coronary Artery Disease	5	10.53%	<0.001
Dependent Functional Status	8	2.35%	0.007
Prolonged Operative Time	34	2.53%	<0.001

^aComparing the number of patients in each category with and without VTE. Bolded values indicate significance at *P* < 0.05.

Table 3. Multivariable Logistic Regression to Determine Predictive Factors in Training Set.

	OR	95% CI	<i>P</i> -value ^a
Elderly (Age > 65)	1.52	1.07—2.17	0.019
Obesity (BMI > 35)	1.56	1.07—2.26	0.019
Coronary artery disease	10.52	2.33—47.6	0.002
Dependent functional status	2.43	1.16—4.97	0.019
Operative time (>304 minutes)	3.27	2.19—4.87	<0.001

^aBolded values indicate significance at *P* < 0.05.

Predictive Model Validation

For the 2,750 patients included in the test dataset, each increase in the predictive model score was associated with a significant and incremental increase in the odds of developing VTE.

Table 4. Odds Ratio for Increased Risk of VTE for Each Score Using the Validation Set.

Score	OR	95% CI	P-Value
1	1.89	1.14—3.12	0.013
2	1.91	1.05—3.45	0.033
3	2.81	1.58—5.01	< 0.001
4	4.85	2.25—10.42	< 0.001
5	5.14	2.31—11.45	< 0.001
6	10.98	4.68—25.77	< 0.001
7	99.63	8.76—1133.08	< 0.001
8*	-	-	-

All odds ratios are in reference to a score of 0.

Bolded values indicate significance at $P < 0.05$.

*There were no patients with a score of 8.

Table 5. Risk of VTE for Each Score.

Score	Risk (%)
0	0.50%
1	0.94%
2	0.95%
3	1.39%
4	2.37%
5	2.52%
6	5.22%
7	33.33%
8*	-

*There were no patients with a score of 8.

Compared to patients with a score 0, patients with a score 1 had 1.9x-odds of VTE development, and those with a score of 7 had 99.6x-odds of VTE ($P < 0.05$ for all; Table 4).

Patients had a 0.50% risk of a VTE with a baseline score of 0, 0.94% risk with a score of 1, 0.95% with a score of 2, 1.39% with a score of 3, 2.37% with a score of 4, 2.52% with a score of 5, 5.22% with a score of 6, and 33.33% with a score of 7 (Table 5). The AUC of the predictive model was 0.709 (95% CI [0.694, 0.724]), which was significantly higher than the AUC of 0.469 (95% CI [0.443, 0.495]) for ASA and the AUC of 0.515 (95% CI [0.496, 0.534]) for CCI (all, $P < 0.001$). The empirical optimal cut-off for our scoring system was determined to be 3 ($J = 0.378$) using Youden's analysis. This cutoff point had a corresponding positive likelihood ratio (LR+) of 1.58 (Table 6). In other words, patients with score of 0, 1, or 2 are considered to be at low risk for VTE based on our model, and those with scores 3 and above were considered to be high risk for VTE.

XGBoost Parameters

For construction of the machine learning model, we initially tested 17 variables for inclusion into the XGBoost model: elderly (age over 65-years-old), sex, smoking status, functional status, diabetes, COPD, CHF, CAD, dyspnea, renal failure, steroid use, transfusion history, bleeding disorder, hypoalbuminemia

Table 6. Sensitivity, Specificity, and Positive Likelihood Ratios for Each Score.

Score	Sensitivity (%)	Specificity (%)	LR+
0	100.00%	0.00%	1.00
1	79.69%	38.75%	1.30
2	50.78%	67.95%	1.58
3	35.94%	82.81%	2.09
4	19.53%	93.95%	3.23
5	12.50%	96.72%	3.81
6	6.25%	99.04%	6.48
7	0.78%	99.99%	52.24
8*	-	-	-

*Metrics unable to be calculated since there were no patients with a score of 8.

(albumin below 3.5 g/dl), anemia (hematocrit below 30%), hyper-bilirubinemia (total bilirubin above 2.5), abnormal INR (INR above 1), and prolonged operative time (over 304 minutes). To prevent overfitting, the 10 most predictive variables were used to build the XGBoost algorithm: BMI, operative time, hematocrit, age, albumin, CAD, gender, smoking status, COPD, and functional status. A feature importance plot demonstrating the relative scaled importance of each continuous and categorical variable used in the XGBoost algorithm is displayed in Figure 1.

XGBoost Validation

The XGBoost model had an AUC of 0.716 (95% CI [0.701, 0.731]), which was significantly greater than the AUCs of ASA and CCI ($P < 0.001$ for all). There was no statistically significant difference between the AUCs of the XGBoost model and the predictive model (0.716 vs. 0.709 respectively, $P > 0.05$).

Discussion

Standardized risk-stratification tools can aid surgeons and peri-operative teams in identifying patients who are at risk of post-operative complications. The aim of our study was to leverage predictive analytics and machine learning in developing novel tools for predicting postoperative VTE following single-level PLF.

To this end, we developed and internally validated 2 novel tools for predicting VTE. First, we present a simple and predictive 5-variable clinical score using age, BMI, history of coronary artery disease, functional status, and operative time to identify patients at risk of VTE following single-level PLF. Then, we developed an XGBoost machine learning algorithm incorporating 10 preoperative variables into a computationally complex, tree-based model. Both the simple scoring tool and complex XGBoost algorithm are significantly more predictive than benchmark ASA and CCI metrics for predicting post-operative VTE. Interestingly, the predictive ability between the 2 models are comparable, with no significant difference in terms of AUC.

With the utilization rate of PLF increasing annually in the United States, there is growing concern surrounding the poor

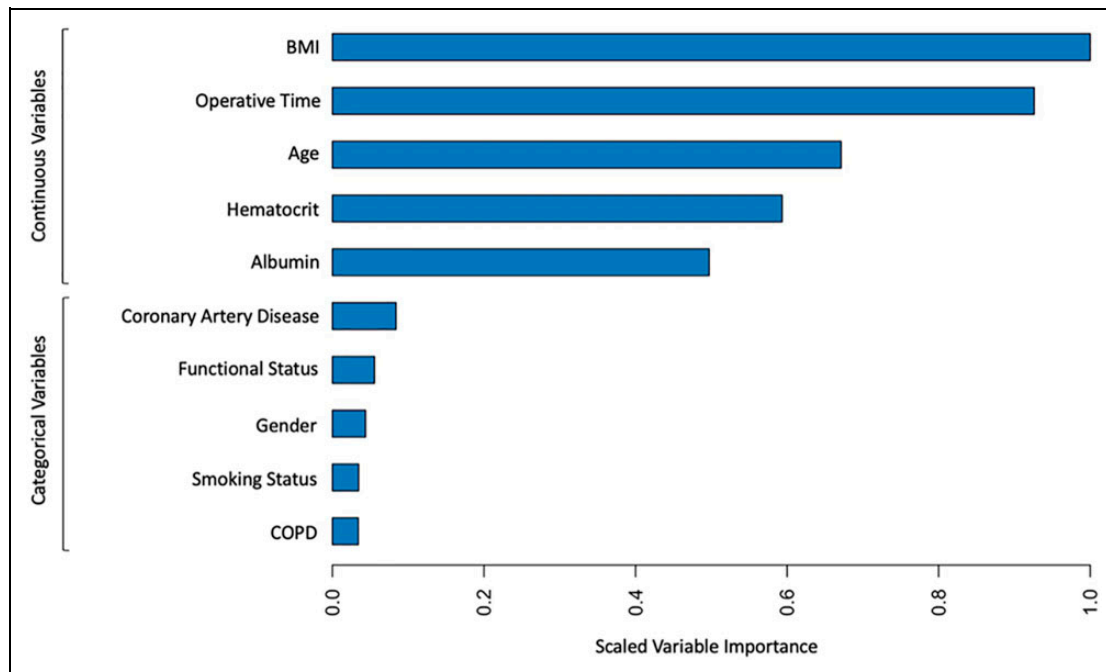


Figure 1. Feature importance plot for XGBoost machine learning model.

outcomes and excess healthcare costs associated with postoperative VTE.⁴ A 2018 review recognized postoperative VTE as the second most common cause of extended length of stay as well as the third most common cause of increased mortality and healthcare costs in hospitalized surgical patients.²⁷ In terms of quality improvement, VTE prevention has been a continued area of focus for enhancing safety measures and quality of surgical care for the past 2 decades.²⁸ Despite numerous reviews and practice guidelines supporting the efficacy of primary thromboprophylaxis in patient cohorts at elevated risk of postoperative VTE, several recent studies cite the underutilization of prophylactic measures.^{27,29,30} Consequently, the rates of postoperative VTE has remained relatively constant over time, while the incidence of other postoperative complications have decreased.^{9,31-33} These data suggest a notable gap between evidence-based recommendations and implementation into clinical practice.

A recent meta-analysis of 26 studies found advanced age, female sex, diabetes, chronic kidney disease, non-ambulatory activity status, high D-dimer level, long operative time, and blood transfusion to be factors associated with VTE development following spinal surgery.⁹ These factors are congruent with those found in our predictive model. Additional studies have reported similar risk factors.³²⁻³⁴ As spine surgery also carries the opposing risk of bleeding complications such as epidural hematoma and associated neurologic injury, surgeons must weigh the risk of over-anticoagulation against the benefits of pharmacologic prophylaxis.^{31,34} In the absence of a standardized stratification system for identifying patients at risk of postoperative VTE, clinical decision-making regarding patients in whom additional chemoprophylaxis is necessary can be challenging.

On AUC analysis, our scoring system proved to be more predictive than benchmark risk-stratification assessments like

ASA and CCI for identifying patients at risk of VTE following PLF. In the present study, the AUCs for both ASA and CCI were around 0.5, suggesting that the predictive ability of these metrics for VTE was not higher than random chance. The proposed predictive model scoring tool is able to build on these legacy metrics, and is directly tailored toward predicting postoperative VTE while retaining the simplicity of commonly used legacy metrics.

Machine learning algorithms such as gradient tree boosting have been increasingly leveraged to aid in clinical and surgical decision-making.^{35,36} Given enormous quantities of data, machine learning offers distinct advantages over traditional analytic methods in terms of increased speed, processing power, and ability to identify complex associations amid substantial statistical noise.²³ The “ensemble” technique of the XGBoost algorithm allows it to build upon iterative tree-based models to converge into a final, highly predictive model, which was demonstrated in its ability to predict VTE in the present study.

XGBoost models have also been successfully applied in other healthcare domains. Within the field of orthopedics, Lu et al and Kumar et al demonstrated the efficacy of XGBoost in preoperative screening and risk stratification of patients likely to have poor clinical outcomes following ACL reconstruction and shoulder arthroplasty, respectively.^{37,38} In the present study, our XGBoost model was also highly predictive of VTE, but it did not outperform our simpler, 5-factor predictive tool despite incorporating twice the amount of user input. Ultimately, our predictive tool is as effective as the XGBoost algorithm, but its simplicity makes it practical in a clinical setting.

Contrary to our initial hypothesis, the computationally complex XGBoost model and the relatively simpler scoring tool

were comparable in terms of predictive ability when applied to our validation dataset. This may suggest that, at present, machine learning models may not necessarily be more predictive than simpler models which are more convenient to utilize in clinical practice. Machine learning models likely require implementation into an existing electronic medical record system with automatic input of the variables through the electronic medical record system to be practically used real-time in a clinical setting. It is important to note, however, that machine learning models may prove to be more predictive than simpler models when incorporating more data points, in terms of both quantity and granularity, than those which are provided in the NSQIP database.

This study has several limitations. First, there are variables associated with a patient's risk of VTE that are not captured in NSQIP data. For instance, we did not have access to pharmacologic data, so we are unable to assess the impact of agents such as preoperative oral contraceptives or perioperative tranexamic acid on VTE. Similarly, we are unable to identify which patients in our cohort received thromboprophylaxis during the perioperative period, which may substantially alter a patient's risk of VTE. A prior study by McLynn et al found that pharmacologic prophylaxis, primarily with unfractionated heparin, after elective spine surgery was not associated with a significant reduction in VTE.³⁴ Second, we are unable to assess a patient's level of postoperative activity, which maybe a risk-factor for VTE. Third, because only 1 dataset was used for model derivation and internal validation, we may have inadvertently overfitted the models. Future studies should seek to externally validate our predictive tools. Finally, we only focused on 1-level fusions, and fusion levels maybe a risk factor for VTE development. Future investigations should examine the potential cost savings provided by predictive analytics models which can accurately identify patients at risk of VTE following spine surgery.

Conclusion

We present a simple, 5-variable predictive model (patient age > 65, obesity grade II or above, coronary artery disease, functional status, and prolonged operative time) to predict patients at risk of VTE following single-level PLF surgery. The predictive model achieved similar accuracy in VTE prediction as a sophisticated machine learning model. Surgeons and perioperative teams may find these tools useful to augment clinical decision-making and perioperative management.

Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Ijezie Ikwuezunma, BS, BA  <https://orcid.org/0000-0002-0775-8506>

Amit Jain, MD  <https://orcid.org/0000-0002-9983-3365>

Supplemental Material

Supplemental material for this article is available online.

References

1. Park MS, Ju Y-S, Moon S-H, et al. Reoperation rates after posterior lumbar spinal fusion surgery according to preoperative diagnoses: a national population-based cohort study. *Clin Neurol Neurosurg.* 2019;184:105408. doi:10.1016/j.clineuro.2019.105408
2. Sheikh SR, Thompson NR, Benzel E, et al. Can we justify it? Trends in the utilization of spinal fusions and associated reimbursement. *Neurosurgery.* 2020;86(2):E193-E202. doi:10.1093/neuros/nyz400
3. Yamaguchi JT, Weiss HK, Garcia RM, et al. Trends in national utilization of posterior lumbar fusion and 30-day reoperation and readmission rates from 2006-2016. *Clin Neurol Neurosurg.* 2020; 199:106310. doi:10.1016/j.clineuro.2020.106310
4. Pannell WC, Savin DD, Scott TP, Wang JC, Daubs MD. Trends in the surgical treatment of lumbar spine disease in the United States. *Spine J.* 2015;15(8):1719-1727. doi:10.1016/j.spinee.2013.10.014
5. Yoshihara H, Yoneoka D. National trends in the surgical treatment for lumbar degenerative disc disease: United States, 2000 to 2009. *Spine J.* 2015;15(2):265-271. doi:10.1016/j.spinee.2014.09.026
6. Wilson LA, Fiasconaro M, Liu J, et al. Trends in comorbidities and complications among patients undergoing inpatient spine surgery. *Spine (Phila Pa 1976).* 2020;45(18):1299-1308. doi:10.1097/BRS.0000000000003280
7. Weber B, Seal A, McGirr J, Fielding K. Case series of elective instrumented posterior lumbar spinal fusions demonstrating a low incidence of venous thromboembolism. *ANZ J Surg.* 2016;86(10): 796-800. doi:10.1111/ans.12702
8. Yoshioka K, Murakami H, Demura S, et al. Comparative study of the prevalence of venous thromboembolism after elective spinal surgery. *Orthopedics.* 2013;36(2):e223-e228. doi:10.3928/01477447-20130122-26
9. Zhang L, Cao H, Chen Y, Jiao G. Risk factors for venous thromboembolism following spinal surgery: a meta-analysis. *Medicine (Baltimore).* 2020;99(29):e20954. doi:10.1097/MD.00000000000020954
10. Phillippe HM. Overview of venous thromboembolism. *Am J Manag Care.* 2017;23(20 suppl):S376-S382.
11. Kraus VB, Blanco FJ, Englund M, Karsdal MA, Lohmander LS. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. *Osteoarthritis Cartilage.* 2015;23(8):1233-1241. doi:10.1016/j.joca.2015.03.036
12. Meneghini RM, Ziemba-Davis M, Ishmael MK, Kuzma AL, Caccavallo P. Safe selection of outpatient joint arthroplasty patients with medical risk stratification: the "Outpatient Arthroplasty Risk Assessment Score." *J Arthroplasty.* 2017;32(8):2325-2331. doi: 10.1016/j.arth.2017.03.004

13. Bohl DD, Idarraga AJ, Holmes GB Jr, Hamid KS, Lin J, Lee S. Validated risk-stratification system for prediction of early adverse events following open reduction and internal fixation of closed ankle fractures. *J Bone Joint Surg Am.* 2019;101(19):1768-1774. doi:10.2106/JBJS.19.00203
14. Wang KY, Suresh KV, Puvanesarajah V, Raad M, Margalit A, Jain A. Using predictive modeling and machine learning to identify patients appropriate for outpatient ACDF. *Spine (Phila Pa 1976).* 2021; 46(10):665-670. doi:10.1097/BRS.0000000000003865
15. Bohl DD, Maltenfort MG, Huang R, Parvizi J, Lieberman JR, Della Valle CJ. Development and validation of a risk stratification system for pulmonary embolism after elective primary total joint arthroplasty. *J Arthroplasty.* 2016;31(9 suppl):187-191. doi:10.1016/j.arth.2016.02.080
16. Ondeck NT, Bohl DD, Bovonratwet P, et al. Discriminative ability of commonly used indices to predict adverse outcomes after poster lumbar fusion: a comparison of demographics, ASA, the modified Charlson comorbidity index, and the modified frailty index. *Spine J.* 2018;18(1):44-52. doi:10.1016/j.spinee.2017.05.028
17. Shultz BN, Ottesen TD, Ondeck NT, et al. Systematic changes in the national surgical quality improvement program database over the years can affect comorbidity indices such as the modified frailty index and modified Charlson comorbidity index for lumbar fusion studies. *Spine (Phila Pa 1976).* 2018;43(11):798-804. doi:10.1097/BRS.0000000000002418
18. Whitmore RG, Stephen JH, Vernick C, et al. ASA grade and Charlson comorbidity index of spinal surgery patients: correlation with complications and societal costs. *Spine J.* 2014;14(1):31-38. doi:10.1016/j.spinee.2013.03.011
19. Shiloach M, Frencher SKJ, Steeger JE, et al. Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. *J Am Coll Surg.* 2010;210(1):6-16. doi:10.1016/j.jamcollsurg.2009.09.031
20. Hustedt JW, Chung A, Bohl DD. Development of a risk stratification scoring system to predict general surgical complications in hand surgery patients. *J Hand Surg Am.* 2018;43(7):641-648.e6. doi:10.1016/j.jhssa.2018.05.001
21. Huang Z, Hu C, Chi C, Jiang Z, Tong Y, Zhao C. An artificial intelligence model for predicting 1-year survival of bone metastases in non-small-cell lung cancer patients based on XGBoost algorithm. *Biomed Res Int.* 2020;2020:3462363. doi:10.1155/2020/3462363
22. Ogunleye A, Wang Q-G. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinforma.* 2020;17(6):2131-2140. doi:10.1109/TCBB.2019.2911071
23. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health.* 2019;22(7):808-815. doi:10.1016/j.jval.2019.02.012
24. Saklad M. Grading of patients for surgical procedures. *Anesthesiology.* 1941;2(3):281-284. doi:10.1097/00000542-194105000-00004
25. Stavem K, Hoel H, Skjaker SA, Haagensen R. Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients. *Clin Epidemiol.* 2017;9:311-320. doi:10.2147/CLEP.S133624
26. De Bruijn B. Revisiting the area under the ROC. *Stud Health Technol Inform.* 2011;169:532-536.
27. Kahn SR, Morrison DR, Dienderé G, et al. Interventions for implementation of thromboprophylaxis in hospitalized patients at risk for venous thromboembolism. *Cochrane Database Syst Rev.* 2018;4(4):CD008201. doi:10.1002/14651858.CD008201.pub3
28. Shojania KG, Duncan BW, McDonald KM, Wachter RM, Markowitz AJ. Making health care safer: a critical analysis of patient safety practices. *Evid Rep Technol Assess (Summ).* 2001;(43):i-x, 1-668.
29. Prandoni P, Sabbion P, Tanduo C, Errigo G, Zanon E, Bernardi E. Prevention of venous thromboembolism in high-risk surgical and medical patients. *Semin Vasc Med.* 2001;1(1):61-70. doi:10.1055/s-2001-14542
30. Kakkar AK, Cohen AT, Tapson VF, et al. Venous thromboembolism risk and prophylaxis in the acute care hospital setting (ENDORSE survey): findings in surgical patients. *Ann Surg.* 2010;251(2):330-338. doi:10.1097/SLA.0b013e3181c0e58f
31. Al-Dujaili TM, Majer CN, Madhoun TE, Kassis SZ, Saleh AA. Deep venous thrombosis in spine surgery patients: incidence and hematoma formation. *Int Surg.* 2012;97(2):150-154. doi:10.9738/CC71.1
32. Wang T, Yang S-D, Huang W-Z, Liu F-Y, Wang H, Ding W-Y. Factors predicting venous thromboembolism after spine surgery. *Medicine (Baltimore).* 2016;95(52):e5776. doi:10.1097/MD.0000000000005776
33. Xin W-Q, Xin Q-Q, Ming H-L, et al. Predictable risk factors of spontaneous venous thromboembolism in patients undergoing spine surgery. *World Neurosurg.* 2019;127:451-463. doi:10.1016/j.wneu.2019.04.126
34. McLynn RP, Diaz-Collado PJ, Ottesen TD, et al. Risk factors and pharmacologic prophylaxis for venous thromboembolism in elective spine surgery. *Spine J.* 2018;18(6):970-978. doi:10.1016/j.spinee.2017.10.013
35. Zhang Z, Zhao Y, Canes A, Steinberg D, Lyashevskaya O. Predictive analytics with gradient boosting in clinical medicine. *Ann Transl Med.* 2019;7(7):152. doi:10.21037/atm.2019.03.29
36. Liu Y, Gu Y, Nguyen JC, et al. Symptom severity classification with gradient tree boosting. *J Biomed Inform.* 2017;75S:S105-S111. doi:10.1016/j.jbi.2017.05.015
37. Lu Y, Forlenza E, Cohn MR, et al. Machine learning can reliably identify patients at risk of overnight hospital admission following anterior cruciate ligament reconstruction [published online October 12, 2020]. *Knee Surg Sports Traumatol Arthrosc.* 2020. doi:10.1007/s00167-020-06321-w
38. Kumar V, Roche C, Overman S, et al. Using machine learning to predict clinical outcomes after shoulder arthroplasty with a minimal feature set. *J Shoulder Elbow Surg.* 2021;30(5):e225-e236. doi:10.1016/j.jse.2020.07.042