# An Artificial Intelligence Approach to Predicting Unplanned Intubation Following Anterior Cervical Discectomy and Fusion

Global Spine Journal 2023, Vol. 13(7) 1849–1855 © The Author(s) 2022 Article reuse guidelines: [sagepub.com/journals-permissions](https://us.sagepub.com/en-us/journals-permissions) DOI: [10.1177/21925682211053593](https://doi.org/10.1177/21925682211053593) [journals.sagepub.com/home/gsj](https://journals.sagepub.com/home/gsj) S Sage

Ashwin Veeramani<sup>l</sup> ©, Andrew S Zhang, MD<sup>2</sup>, Amy Z. Blackburn, BA BM<sup>2</sup>, Christine M. Etzel, ScB<sup>2</sup> <sup>®</sup>, Kevin J. DiSilvestro, MD<sup>2</sup>, Christopher L. McDonald, MD<sup>2</sup>, and Alan H. Daniels, MD<sup>2</sup>

# Abstract

Study Design: Level III retrospective database study.

Objectives: The purpose of this study is to determine if machine learning algorithms are effective in predicting unplanned intubation following anterior cervical discectomy and fusion (ACDF).

Methods: The National Surgical Quality Initiative Program (NSQIP) was queried to select patients who had undergone ACDF. Machine learning analysis was conducted in Python and multivariate regression analysis was conducted in R. C-Statistics area under the curve (AUC) and prediction accuracy were used to measure the classifier's effectiveness in distinguishing cases.

Results: In total, 54 502 patients met the study criteria. Of these patients, .51% underwent an unplanned re-intubation. Machine learning algorithms accurately classified between 72%-100% of the test cases with AUC values of between .52-.77. Multivariable regression indicated that the number of levels fused, male sex, COPD, American Society of Anesthesiologists (ASA) > 2, increased operating time, Age > 65, pre-operative weight loss, dialysis, and disseminated cancer were associated with increased risk of unplanned intubation.

Conclusions: The models presented here achieved high accuracy in predicting risk factors for re-intubation following ACDF surgery. Machine learning analysis may be useful in identifying patients who are at a higher risk of unplanned post-operative reintubation and their treatment plans can be modified to prophylactically prevent respiratory compromise and consequently unplanned re-intubation.

#### Keywords

anterior cervical discectomy and fusion surgery, unplanned intubation, complications

# Introduction

Respiratory compromise is a potentially devastating complication that is well recognized in anterior cervical discectomy and fusion (ACDF) surgery. Current research has identified comorbidities and demographic risk factors associated with re-intubation following these surgeries. Coagulopathies, obesity, rheumatoid arthritis, and chronic obstructive pulmonary disease (COPD) have been linked to unplanned re-intubation among cervical spine surgery pa-tients.<sup>[1](#page-5-0)</sup> Type II diabetes, operating time longer than 5 hours, blood loss greater than 300 mL, and being female have also been associated with increased risk of re-intubation after  $ACDF<sup>2,3</sup>$  $ACDF<sup>2,3</sup>$  $ACDF<sup>2,3</sup>$  $ACDF<sup>2,3</sup>$  While much research has been conducted into

<sup>1</sup>Division of Applied Mathematics, Brown University, Providence, RI, USA <sup>2</sup>Department of Orthopedic Surgery, Rhode Island Hospital, Warren Alpert Medical School of Brown University, Providence, RI, USA

#### Corresponding Author:

Alan Daniels, MD, Department of Orthopedic Surgery, Rhode Island Hospital, Warren Alpert Medical School of Brown University, 1Kettle Point Ave, East Providence, RI 02914, USA. Email: [alandanielsmd@gmail.com](mailto:alandanielsmd@gmail.com)



Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License [\(https://creativecommons.org/licenses/by-nc-nd/4.0/\)](https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages [\(https://us.sagepub.com/en-us/nam/open-access-at-sage\)](https://us.sagepub.com/en-us/nam/open-access-at-sage).

recognizing these risk factors, no research has been performed examining the value of machine learning and its ability to predict unplanned intubation pre-operatively in patients undergoing ACDF.

Machine Learning systems have been previously used in several clinical research areas and proven effective in predicting patient outcomes. Supervised Machine Learning algorithms such as logistic regressions and decision trees have been shown to accurately predict post-operative outcomes such as readmission and blood transfusion needs.<sup>[4,](#page-5-3)[5](#page-5-4)</sup> In orthopedics and spine surgery, machine learning algorithms such as the random forest and decision tree algorithms have proven useful in predicting length of stay in shoulder arthroplasty patients and blood transfusion after spine fusion surgery for adult spinal deformity  $(ASD).$ <sup>[5,](#page-5-4)[6](#page-5-5)</sup>

The aim of this investigation is to utilize a set of rule-based machine learning processes to determine which patients undergoing ACDF surgery will also endure an unplanned intubation. We hypothesize that the LR, DT, RF, GB, XGB, and NN algorithms will be systematically effective models that can predict the risk of an unplanned intubation following ACDF surgery.

## **Methods**

Data from the American College of Surgeon's National Surgical Quality Initiative Program (NSQIP) was queried to select patients who underwent ACDF surgery between 2010 and 2018. As the data was already collected for general research use and made publicly available before the study was considered and completed, no informed consent was required. Moreover, no institutional review board approval was necessary as the patient data was deidentified. Patients who had undergone ACDF non-electively were excluded from the analysis. Demographic and comorbidity data were collected from the NSQIP database, namely: obesity, number of spine levels fused, diabetes, operating time longer than 250 minutes, age older than 65 years old, gender, ascites, renal failure, race, chronic obstructive pulmonary disorder (COPD), congestive heart failure (CHF), American Society of Anesthesiologists (ASA) Class, history of steroid use, disseminated cancer, greater than 10 pounds weight loss before surgery, and history of bleeding disorders, hypertension, albumin levels (AL-BUM), SGOT levels (SGOT), alkaline phosphatase (ALKPH), white blood cell count (WBC), hematocrit levels (HCT), platelet counts (PLATE), partial thromboplastin time (PTT), international normalized ration (INR), creatinine levels (CREAT), sodium levels (SODM), and blood urea nitrogen levels (BUN). Patients who were missing one or more categorical variables were excluded from the dataset. Patients with the remaining missing quantitative variables were imputed with the nearest neighbor algorithm that utilizes correlations among patients to estimate the missing value of a patient characteristic.<sup>[7](#page-5-6)</sup>

A retrospective case control study was conducted. Six machine learning algorithms will be tested: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), and a Neural Network (NN). DT algorithms utilize an eponymous process to classify test cases into one of two cases. RF algorithms utilize a collection of decision trees to help classify data into binary outcomes. The decision tree process follows an algorithm known as the Gini Index. The Gini index analyzes the distribution of data and frequency of values to design a set of decision trees with a high prediction accuracy. $7-10$  $7-10$  $7-10$  GB and XGB algorithms construct highly complex decision trees that are more nuanced than DT and RF algo-rithms to classify patients.<sup>[11,](#page-5-8)[12](#page-5-9)</sup> Finally, NN algorithms utilize a set of nodes arranged similar to neural connections in the nervous system to predict whether a patient will be readmitted or not[.12](#page-5-9)-[14](#page-5-10) LR, DT, RF, GB, XGB, and NN algorithmic classification analysis was conducted in Python's SciKit Machine Learning Package. The first group consisted of patients who did not undergo unplanned intubation, while the second group consisted of patients that underwent unplanned intubation.<sup>[15](#page-5-11)</sup> The data was split into training and testing groups where 70% of the data was allocated to train the random forest algorithm, while the remaining 30% was used to test the validity of the machine learning models.

Receiver operating curves (ROCs), prediction accuracy, and Brier scores were calculated to measure the classifier's effectiveness in distinguishing cases. ROCs are useful for calculating C-Statistics, also known as area under the curve  $(AUC)$ , through trapezoidal integration.<sup>[16](#page-5-12)</sup> ROC curves show the trade-off between the true positive and true negative identification capacities. Curves that inflect near the top left of graph are highly effective in distinguishing between the test cases as there is little trade-off in the true positive and true negative identification ability of the algorithm.[17](#page-5-13) Feature importance was determined in order to identify the variables most responsible for the machine learning's distinction process. Prediction accuracy was useful for determining how accurate the model was pre-dicting all the test cases as a holistic metric.<sup>[18](#page-5-14)</sup> Brier scores are a determination of the probabilistic accuracy for a given algorithm. Brier scores close to 0 indicate strong probabilistic accuracy, while Brier scores close to 1 indicate poor probabilistic accuracy.<sup>[19](#page-5-15)-[21](#page-5-16)</sup>

Multivariate logistic regression analysis was conducted in R (version 3.6.3). Demographic and comorbidity data were used to predict risk factors for unplanned intubation. Odds ratios along with 95% confidence intervals and associated Pvalues were extracted from the regression data. The significance level was set at .05.

## **Results**

In total, 62 934 patients met the study criteria. Patients that underwent emergent surgery or were missing one or more

Table 1. Demographics and Comorbidity Information for the Patient Sample.

	<b>Total Group</b>	No Intubation	Intubation	P-value
BMI $(kg/m2)$	30.36	30.36	29.38	.0274
<b>Diabetes</b>	8615 (15.92%)	8549 (15.87%)	66 (27.73%)	< .0001
Operating time (minutes)	127.31	127.08	178.58	< 0.001
Smoking history	14 563 (26.92%)	14 488 (26.9%)	75 (31.51%)	.1096
Age (years)	54.71	54.68	63.37	< 0.001
% Male	26 816 (49.57%)	26 660 (49.5%)	82 (34.45%)	< .0001
% Caucasian	43 427 (80.31%)	43 242 (80.33%)	185 (77.73%)	.3141
<b>CHF</b>	$107$ (.2%)	104 (.19%)	3(1.26%)	.0002
<b>COPD</b>	2438 (4.51%)	2402 (4.46%)	36 (15.13%)	< 0001
<b>Dialysis</b>	94 (.17%)	89 (.17%)	5(2.1%)	< 0001
ASA > 2	23 199 (42.88%)	23 023 (42.74%)	176 (73.95%)	< .0001
Steroid use	1760 (3.25%)	1747 (3.24%)	13(5.46%)	.054
Dis. cancer	72 (.13%)	68 (.13%)	4(1.68%)	< .0001
Weight loss	76 (.14%)	72 (.13%)	4(1.68%)	< 0001
Bleeding disorder	565 (1.04%)	556 (1.03%)	9(3.78%)	< 0001
Hypertension	24 678 (45.61%)	24 519 (45.52%)	79 (33.19%)	1000.
Album (g/dL)	4.14	4.14	4.01	< .0001
SGOT (u/sL)	25.35	25.35	25.93	.5991
ALKPH (U/L)	77.68	77.64	87.04	< .0001
WBC $(10^9/L)$	7.46	7.45	7.68	.2034
<b>HCT (%)</b>	41.74	41.75	40.86	.001
PLATE $(10^9/L)$	248.80	248.87	233.35	.0003
PTT (seconds)	29.19	29.19	29.76	.0399
INR (units)	1.01	1.01	1.02	.2842
CREAT (mg/dL)	.91	.91	1.10	< 0001
SODM (mEq/L)	139.50	139.50	138.77	< 0.001
BUN (mmol/L)	15.42	15.40	19.19	< 0001
Levels	1.72	1.72	2.02	< 0001
Total	54 102	53 864	238	

<span id="page-2-0"></span>Variable Key: body mass index (BMI) > 30 kg/m<sup>2</sup> (obesity), diabetes history (diabetes), operating time (Op Time), smoking history (Smoke), age, sex, race, congestive heart failure (CHF), chronic obstructive pulmonary disorder (COPD), dialysis history (Dialysis), American Society of Anesthesiologist Class (ASA), steroid use history (Steroid), presence of disseminated cancer (Dis. Cancer), pre-operative weight loss greater than 10 pounds (weight loss), bleeding disorder history (Bleed Disorder), hypertension history (hypertension), albumin levels (Album), aspartate aminotransferase test (SGOT), alkaline phosphatase test (ALKPH), white blood cell count (WBC), hematocrit levels (HCT), platelet count (PLATE), partial thromboplastin time (PTT), international normalized ratio (INR), creatinine levels (CREAT), sodium levels (SODM), blood urea nitrogen (BUN), and number of spine levels fused (levels).

Table 2. Performance of the Machine Learning Algorithms in the Study.

	Prediction Accuracy, %	<b>AUC</b>	<b>Brier Score</b>
LR.	71.7	.766	18.
DT	99.2	.520	.008
<b>RF</b>	99.6	.678	.004
GB	99.5	.737	.005
<b>XGB</b>	99.6	.713	.004
<b>NN</b>	99.6	.701	.004

<span id="page-2-1"></span>Key: LR, Logistic Regression; DT, Decision Tree; RF, Random Forest; GB, Gradient Boosting; XGB, Extreme Gradient Boosting; NN, Neural Network.

categorical variables were excluded from the dataset. Overall, 54 502 (86.7%) were ultimately included, and .51% (n = 238) of these patients underwent a post-operative unplanned reintubation. A demographics table is shown in [Table 1.](#page-2-0) There was an average of 4.5 days between surgery and unplanned intubation and a median of 2 days (25th Percentile: 0 days, 75th Percentile: 6 days).

The machine learning algorithms were able to accurately classify between 71.7% and 99.6% of the test cases with AUC values of between .52 and .77 and Brier scores between .004 and .18. The full results of the machine learning analysis are shown in [Table 2.](#page-2-1) The ROCs of the algorithms are shown in [Figure 1](#page-3-0). The calibration curves that visualize the Brier scores are shown in [Figure 2.](#page-3-1)

The multivariate regression identified the factors linked to unplanned intubation as the number of levels fused, male sex, COPD,  $ASA > 2$ , increased operating time, Age  $> 65$ , pre-operative weight loss, dialysis, and disseminated cancer. The full results of the multivariate regression are summarized in [Table 3](#page-4-0) and visualized in the forest plot in [Figure 3](#page-4-1).

ROC AUC  $1.0$  $0.8$ Sensitivity<br>0.4 LR DT **RF**  $0.2$  $GB$ **XGB** NN  $0.0$  $0.0$  $0.2$  $0.4$  $0.6$  $0.8$  $1.0$ 1-Specificity

<span id="page-3-0"></span>Figure 1. ROC Curve visualizing the AUC scores for each of the algorithms.



<span id="page-3-1"></span>Figure 2. Calibration Curve visualizing the Brier scores for each of the algorithms.

#### **Discussion**

While ACDF surgery is a common surgical procedure, one of the most dangerous complications of ACDF is respiratory failure, ultimately requiring an unplanned re-intubation. Several research studies have been dedicated to identifying these risk factors for unplanned intubations, which highlight the need for a useful model to provide a risk assessment of whether a patient that is planned to proceed with ACDF will experience an unplanned intubation.

This study presents a set of machine learning algorithms that can pre-operatively predict whether a patient undergoing an ACDF will require re-intubation. The machine learning algorithms were able to predict between 71.7% and 99.6% of the test cases accurately with AUC values of between .52 and .77. Based upon prediction accuracy, AUC, and Brier scores, the Gradient Boosting algorithm was found to have performed

the best among all the algorithms tested. As such, an online web calculator for predicting unplanned intubation was created utilizing the Gradient Boosting algorithm: [https://](https://acdfintubation.herokuapp.com/) acdfi[ntubation.herokuapp.com/.](https://acdfintubation.herokuapp.com/)

A multivariate logistic regression identified the number of levels fused, male sex, COPD, ASA > 2, increased operating time, age > 65, pre-operative weight loss, dialysis, smoking, and disseminated cancer as variables that are significantly correlated to unplanned re-intubation. These risk factors agree with the trends in current literature; however, we found that males were more at risk for intubation than females, contrary to what Heyer et al reported. $1-3$  $1-3$ 

The results of the paper presented here indicate that machine learning algorithms show much promise for the prediction of unplanned intubation after surgery and may aid in patient selection for ACDF surgery. Patients who are at risk of unplanned intubation can be better informed of the risks of surgery and providers can better target patient care plans that minimize the risk for an unplanned intubation after ACDF surgery. $21$  This paper demonstrates the efficacy and importance of machine learning integration into clinical settings to predict post-operative spine surgery complications.<sup>[22](#page-6-0)</sup>

The measures of the machine learning models presented here are consistent with the current literature. No research has been conducted specifically on machine learning algorithms to predict unplanned intubation in ACDF patients; however, a few studies have analyzed the accuracy of machine learning algorithms in predicting post-operative complications of spine and shoulder procedures.[5](#page-5-4)[,6,](#page-5-5)[23](#page-6-1)-[25](#page-6-2) Kalagara et al and Durand et al predicted post-operative blood transfusion and readmission using similar models after lumbar laminectomies and fusions. Their results yielded AUC values of .79 and .85 for post-operative blood transfusion and AUC values of .81 and .69 for readmission, respectively. $4,5,26$  $4,5,26$  $4,5,26$  The AUC and prediction accuracies for both studies are comparable to the AUC and prediction accuracies of the algorithms in this study, indicating test cases could accurately be classified. A random forest model to predict post-operative LOS after Total Shoulder Arthroplasty (TSA) with an AUC of .77 was developed in Biron et al.'s study, which is in the upper ranges of the AUC values presented in this paper.<sup>[6](#page-5-5)</sup> Furthermore, Hsieh et al analyzed several machine learning algorithms and their capacity to predict unplanned extubation in ICU patients. The algorithms detailed in the study had AUC values of between .58 and .91. $^{25}$  $^{25}$  $^{25}$  The model presented in this paper falls in this range as well.

The multivariate regression analysis results are in line with current literature. A study on airway management following ACDF conducted by Kim et al concurred with the study's findings that increased operating times were linked to higher risk of unplanned re-intubation following the procedure.<sup>[27](#page-6-4)</sup> Moreover, a NSQIP analysis of re-intubation risk following ACDF undergone by Wilson et al confirmed the study's findings that male sex, elderly age, and comorbidity burden

	Odds Ratio	2.5% Lower Bound	97.5% Upper Bound	P-Value
Obesity	.71	.54	.93	.01
<b>Diabetes</b>	.80	.59	1.08	.10
Elevated op. time	2.44	1.71	3.48	< 0.01
Smoke	1.40	1.04	1.89	.03
Old age	2.53	1.90	3.36	< 0.01
Male	1.71	1.30	2.24	.00
Caucasian	.90	.66	1.24	.52
<b>CHF</b>	2.15	.63	7.34	.22
<b>COPD</b>	2.08	1.42	3.04	< 0.01
Dialysis	4.84	1.86	12.62	5.001
ASA >2	2.32	1.68	3.20	< 0.01
Steroid use	1.17	.66	2.07	.60
Disseminated cancer	5.36	1.82	15.76	.002
Weight loss	4.68	1.57	14.00	.005
Bleed disorder	1.91	.96	3.78	.06
Hypertension	1.30	.96	1.76	.09
Levels fused	1.22	1.10	1.37	.0004

Table 3. Results of the Multivariate Regression. Odds Ratios, 95% Confidence Intervals, and P-values are provided.

<span id="page-4-0"></span>Key: CHF, congestive heart failure; COPD, chronic obstructive pulmonary disorder; ASA, American Society of Anesthesiologists Class.



<span id="page-4-1"></span>Figure 3. A Forest Plot visualizing the results of the multivariate regression identifying risk factors for unplanned intubation following ACDF surgery.

(dialysis, disseminated cancer, etc.) were associated with a higher risk of unplanned re-intubation. $^{28}$ 

There are several potential limitations inherent to this study. First, the sample of patients undergoing ACDF may not have been accurately representative of all the patients who undergo ACDF. The NSQIP dataset is reliant on reporting from participating hospitals. Thus, the sample 3of ACDF patients between 2010 and 2018 will be overrepresented with patients from hospitals that have the infrastructure to maintain NSQIP reporting standards[.29](#page-6-6) Furthermore, it is nearly impossible to quantify every factor that may pre-dispose to unplanned intubation in anterior cervical spinal surgery and confounding factors may exist among the sample under study. Therefore, the model may not be perfectly accurate since it is based solely on variables that can be quantified as continuous or categorical.

# Conclusion

The models presented here achieved a high accuracy for reintubation following ACDF surgery. The results of this paper prompt research into methodologies for introducing and assessing machine learning analysis into the electronic medical record system for use in clinical settings.

#### Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Alan Daniels MD receives consulting fees from Stryker, Orthofix, Spineart, and EOS. The remaining authors have no conflicts to report.

#### Funding

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

## ORCID iD

Ashwin Veeramani **ID** <https://orcid.org/0000-0002-4429-0834> Christine M. Etzel ScB <https://orcid.org/0000-0003-4302-9498>

#### **References**

- <span id="page-5-0"></span>1. Nandyala SV, Marquez-Lara A, Park DK, et al. Incidence, risk factors, and outcomes of postoperative airway management after cervical spine surgery. Spine. 2014;39(9):E557-E563. doi[:10.](https://doi.org/10.1097/BRS.0000000000000227) [1097/BRS.0000000000000227.](https://doi.org/10.1097/BRS.0000000000000227)
- <span id="page-5-1"></span>2. Kim M, Rhim SC, Roh SW, Jeon SR. Analysis of the risk factors associated with prolonged intubation or reintubation after anterior cervical spine surgery. J Kor Med Sci. 2018;33(17). doi: [10.3346/jkms.2018.33.e77](https://doi.org/10.3346/jkms.2018.33.e77).
- <span id="page-5-2"></span>3. Heyer JH, Cao N, Amdur RL, Rao RR. Postoperative complications following orthopedic spine surgery: Is there a difference between men and women? Int J Spine Surg. 2019;13(2): 125-131. doi[:10.14444/6017](https://doi.org/10.14444/6017).
- <span id="page-5-3"></span>4. Kalagara S, Eltorai AEM, Durand WM, DePasse JM, Daniels AH, Machine learning modeling for predicting hospital readmission following lumbar laminectomy. J Neurosurg Spine. 2019, 30(3):344-352. doi: [10.3171/](https://doi.org/10.3171/2018.8.SPINE1869) [2018.8.SPINE1869.](https://doi.org/10.3171/2018.8.SPINE1869)
- <span id="page-5-4"></span>5. Durand WM, DePasse JM, Daniels AH. Predictive modeling for blood transfusion after adult spinal deformity surgery. Spine. 2018;43(15):1058-1066. doi[:10.1097/BRS.0000000000002515.](https://doi.org/10.1097/BRS.0000000000002515)
- <span id="page-5-5"></span>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. 2019;28:e580-e585. doi: [10.5435/jaaos-d-19-00395.](https://doi.org/10.5435/jaaos-d-19-00395)

- <span id="page-5-6"></span>7. Beretta L, Santaniello A. Nearest neighbor imputation algorithms: A critical evaluation. BMC Med Inf Decis Making. 2016; 16(suppl 3). doi:[10.1186/s12911-016-0318-z.](https://doi.org/10.1186/s12911-016-0318-z)
- 8. Menze BH, Kelm BM, Masuch R, et al, A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC Bioinf, 2009; 10(1):213, doi: [10.1186/1471-](https://doi.org/10.1186/1471-2105-10-213) [2105-10-213](https://doi.org/10.1186/1471-2105-10-213).
- 9. Wongvibulsin S, Wu KC, Zeger SL. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. BMC Med Res Methodol. 2019;20(1):1. doi:[10.1186/s12874-019-0863-0.](https://doi.org/10.1186/s12874-019-0863-0)
- <span id="page-5-7"></span>10. Touw WG, Bayjanov JR, Overmars L, et al, Data mining in the life sciences with random forest: a walk in the park or lost in the jungle? Briefings Bioinf, 2013;14(3):315-326. doi: [10.1093/bib/bbs034](https://doi.org/10.1093/bib/bbs034).
- <span id="page-5-8"></span>11. Denisko D, Hoffman MM. Classification and interaction in random forests. Proc Natl Acad Sci Unit States Am. 2018; 115(8):1690-1692. doi:[10.1073/pnas.1800256115.](https://doi.org/10.1073/pnas.1800256115)
- <span id="page-5-9"></span>12. Zhang Z, Zhao Y, Zhao Y, Canes A, Steinberg D, Lyashevska O. Predictive analytics with gradient boosting in clinical medicine. Ann Transl Med. 2019;7(7):152. doi[:10.21037/atm.2019.03.29.](https://doi.org/10.21037/atm.2019.03.29)
- 13. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-510. doi[:10.1038/s41568-018-0016-5](https://doi.org/10.1038/s41568-018-0016-5).
- <span id="page-5-10"></span>14. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019;19(1):64. doi[:10.1186/s12874-019-0681-4.](https://doi.org/10.1186/s12874-019-0681-4)
- <span id="page-5-11"></span>15. Varoquaux G, Buitinck L, Louppe G, Grisel O, Pedregosa F, Mueller A. Scikit-learn. GetMobile: Mobile Comput Commun. 2015;19(1):29-33. doi[:10.1145/2786984.2786995](https://doi.org/10.1145/2786984.2786995).
- <span id="page-5-12"></span>16. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Casp J Int Med. 2013;4(2):627-635.
- <span id="page-5-13"></span>17. Obuchowski NA, Bullen JA. Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. Phys Med Biol. 2018;63(7):07TR01. doi[:10.](https://doi.org/10.1088/1361-6560/aab4b1) [1088/1361-6560/aab4b1](https://doi.org/10.1088/1361-6560/aab4b1).
- <span id="page-5-14"></span>18. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inf Decis Making. 2019;19(1). doi[:10.1186/s12911-](https://doi.org/10.1186/s12911-019-1004-8) [019-1004-8.](https://doi.org/10.1186/s12911-019-1004-8)
- <span id="page-5-15"></span>19. Rufibach K. Use of Brier score to assess binary predictions. J clin Epidemiol. 2010;63(8):938-939. doi[:10.1016/j.jclinepi.](https://doi.org/10.1016/j.jclinepi.2009.11.009) [2009.11.009.](https://doi.org/10.1016/j.jclinepi.2009.11.009)
- 20. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. J Am Coll Surg. 2013;217(5): 833-842. doi[:10.1016/j.jamcollsurg.2013.07.385](https://doi.org/10.1016/j.jamcollsurg.2013.07.385).
- <span id="page-5-16"></span>21. Milgrom DP, Njoku VC, Fecher AM, Kilbane EM, Pitt HA. Unplanned intubation: When and why does this deadly complication occur? Surgery. 2013;154(2):376-383. doi[:10.1016/j.](https://doi.org/10.1016/j.surg.2013.05.006) [surg.2013.05.006](https://doi.org/10.1016/j.surg.2013.05.006).
- <span id="page-6-0"></span>22. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Fut Healthcare J. 2019;6(2):94-98. doi:[10.7861/](https://doi.org/10.7861/futurehosp.6-2-94) [futurehosp.6-2-94](https://doi.org/10.7861/futurehosp.6-2-94).
- <span id="page-6-1"></span>23. Galbusera F, Casaroli G, Bassani T. Artificial intelligence and machine learning in spine research. JOR Spine. 2019;2(1): e1044. doi[:10.1002/jsp2.1044.](https://doi.org/10.1002/jsp2.1044)
- 24. Malik AT, Khan SN. Predictive modeling in spine surgery. Ann Transl Med. 2019;7(S5):S173. doi:[10.21037/atm.2019.07.99.](https://doi.org/10.21037/atm.2019.07.99)
- <span id="page-6-2"></span>25. Hsieh MH, Hsieh MJ, Chen C-M, Hsieh C-C, Chao C-M, Lai C-C. Comparison of machine learning models for the prediction of mortality of patients with unplanned extubation in intensive care units. Sci Rep. 2018;8(1):1-7. doi[:10.1038/s41598-018-35582-2.](https://doi.org/10.1038/s41598-018-35582-2)
- <span id="page-6-3"></span>26. Kalagara S, Eltorai AEM, Durand WM, DePasse JM, Daniels AH, Machine learning modeling for predicting hospital

readmission following lumbar laminectomy, J Neurosurg Spine, 30, 3, 344–352, Mar. 2019, doi: [10.3171/2018.8.SPINE1869.](https://doi.org/10.3171/2018.8.SPINE1869)

- <span id="page-6-4"></span>27. Kim M, Rhim SC, Roh SW, Jeon SR. Analysis of the Risk Factors Associated with Prolonged Intubation or Reintubation after Anterior Cervical Spine Surgery. J Korean Med Sci. 2018; 33(17):e77. doi:[10.3346/jkms.2018.33.e77](https://doi.org/10.3346/jkms.2018.33.e77).
- <span id="page-6-5"></span>28. Wilson LA, Zubizarreta N, Bekeris J, et al. Risk factors for reintubation after anterior cervical discectomy and fusion surgery: evaluation of three observational data sets. Can J Anest. 2019;67(1):42-56. doi:[10.1007/S12630-019-01492-8.](https://doi.org/10.1007/S12630-019-01492-8)
- <span id="page-6-6"></span>29. Ismael HN, Cox S, Cooper A, Narula N, Aloia T. The morbidity and mortality of hepaticojejunostomies for complex bile duct injuries: a multi-institutional analysis of risk factors and outcomes using NSQIP. HPB. 2017;19(4):352-358. doi[:10.1016/j.hpb.2016.12.004.](https://doi.org/10.1016/j.hpb.2016.12.004)