The Journal of Foot & Ankle Surgery 58 (2019) 410–416

Contents lists available at ScienceDirect

The Journal of Foot & Ankle Surgery

journal homepage: www.jfas.org

Original Research

Machine Learning Accurately Predicts Short-Term Outcomes Following Open Reduction and Internal Fixation of Ankle Fractures

Robert K. Merrill, MD[1], Rocco M. Ferrandino, MD[2], Ryan Hoffman, MD[1], Gene W. Shaffer, MD[3], Anthony Ndu, MD[3]

[1] Orthopedic Surgery Resident, Department of Orthopedic Surgery, Albert Einstein Medical Center, Philadelphia, PA
[2] Ear Nose and Throat Resident, Icahn School of Medicine at Mount Sinai, New York, NY
[3] Orthopedic Surgeon, Department of Orthopedic Surgery, Albert Einstein Medical Center, Philadelphia, PA

ARTICLE INFO

Level of Clinical Evidence: 3

Keywords:
ankle fracture
ankle ORIF
gradient boosting
logistic regression
machine learning
readmissions

ABSTRACT

Ankle fractures are common orthopedic injuries with favorable outcomes when managed with open reduction and internal fixation (ORIF). Several patient-related risk factors may contribute to poor short-term outcomes, and machine learning may be a valuable tool for predicting outcomes. The objective of this study was to evaluate machine-learning algorithms for accurately predicting short-term outcomes after ORIF for ankle fractures. The Nationwide Inpatient Sample and Nationwide Readmissions Database were queried for adult patients ≥ 18 years old who underwent ORIF of an ankle fracture during 2013 or 2014. Morbidity and mortality, length of stay > 3 days, and 30-day all-cause readmission were the outcomes of interest. Two machine-learning models were created to identify patient and hospital characteristics associated with the 3 outcomes. The machine learning models were evaluated using confusion matrices and receiver operating characteristic area under the curve values. A total of 16,501 cases were drawn from the Nationwide Inpatient Sample and used to assess morbidity and mortality and length of stay > 3 days, and 33,504 cases were drawn from the Nationwide Readmissions Database to assess 30-day readmission. Older age, Medicaid, Medicare, deficiency anemia, congestive heart failure, chronic lung disease, diabetes, hypertension, and renal failure were the variables associated with a statistically significant increased risk of developing all 3 adverse events. Logistic regression and gradient boosting had similar area under the curve values for each outcome, but gradient boosting was more accurate and more specific for predicting each outcome. Our results suggest that several comorbidities may be associated with adverse short-term outcomes after ORIF of ankle fractures, and that machine learning can accurately predict these outcomes.

Ankle fractures are among the most common orthopedic trauma injuries in the United States and are the most common fractures seen in the foot and ankle (1–4). Nonoperative management is effective for the majority of ankle fractures, but a large percentage require open reduction and internal fixation (ORIF) (5–7). Outcomes following ORIF are generally favorable, but a number of studies have identified risk factors for poor functional outcome and short-term complications following surgery (8–12).

Machine learning is a powerful method for predicting outcomes that has been used throughout a number of clinical specialties, ranging from oncology to orthopedic surgery (13–17). Machine learning involves providing data to a computational algorithm for training. The training process allows the algorithm to learn how certain input variables are related to an outcome variable and to predict outcomes on new datasets the algorithm has not yet encountered. Once trained, the algorithm can be applied to predict an outcome (17). A variety of different machine learning algorithms exist, with no clear indication of which to use in a given situation.

In the orthopedic literature, machine learning has been used in a wide array of applications that have demonstrated it to be valuable in the clinical setting. Studies have used machine-learning technology to identify pathologic gait patterns, progression of osteoarthritis, and outcomes following lumbar spine fusion (18–22). No studies, however, have investigated machine learning as a tool for prediction of outcomes following ankle fracture surgery. Machine learning may be of benefit to the foot and ankle surgeon for identifying high-risk patients and potentially poor outcomes. The objective of the present study was therefore to evaluate the accuracy of 2 different machine-learning algorithms for predicting short-term outcomes following ORIF of ankle fractures.

Materials and Methods

Patient Sample

This cross-sectional study used data from the 2013 and 2014 Nationwide Inpatient Sample (NIS) and Nationwide Readmissions Database (NRD), both part of the Healthcare

Financial Disclosure: None reported.
Conflict of Interest: None reported.
Address correspondence to: Robert K. Merrill, MD, Albert Einstein Medical Center, 5501 Old York Road, Philadelphia, PA 19141.
E-mail addresses: merrillr@einstein.edu, robmerrill7@gmail.com (R.K. Merrill).

1067-2516/$ - see front matter
https://doi.org/10.1053/j.jfas.2018.09.004

Cost and Utilization Project, Agency for Healthcare Research and Quality (23,24). *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnosis codes were used to identify all patients who were hospitalized with an ankle fracture diagnosis (824.0 to 824.9). Procedure codes (79.30, 79.36, and 79.39) were then used to select all ankle fracture patients who underwent ORIF. The NIS is made up of a 20% sample of all national hospital admissions, and the NRD is made up of a 50% sample of all national hospital admissions, so we expected our sample sizes to be different between the 2 databases.

Inclusion and Exclusion Criteria

Only adults ≥18 years old were included in the study. Unspecified fractures of the ankle were excluded. Patients with a diagnosis related to an infected or failed orthopedic implant were excluded from the study, as these were likely revision cases. Patients with polytrauma, which included a diagnosis code of trauma to another region of the body, in addition to the ankle fracture, were excluded. For the NRD cases only, patients admitted during the month of December were excluded to allow enough follow-up time to capture a 30-day readmission.

Variables and Outcomes

Three short-term outcomes of interest were investigated: inpatient morbidity and mortality, length of stay (LOS) >3 days, and 30-day all-cause readmission. Morbidity and mortality and LOS were derived from the NIS, whereas 30-day all-cause readmissions were derived from the NRD. Patients are not linked between the 2 databases.

Morbidity and mortality was defined by intraoperative complications, postoperative complications that occurred during the same admission, and death that occurred during the same admission. Intraoperative complications included accidental laceration or puncture to the patient and hemorrhage complicating the procedure. Postoperative complications involved wound complications, infection, and complications related to the gastrointestinal, urinary, respiratory, and cardiovascular organ systems.

The NIS and NRD include variables to describe various hospital and patient characteristics. Insurance status was grouped into categories that included Medicare, Medicaid, private insurance, self-pay, no charge, and other. The self-pay, no charge, and other insurance categories were combined into a single categorical variable defined as “other insurance.” For the teaching hospital variable, rural hospitals were considered nonteaching, as very few rural teaching hospitals are part of the NIS and NRD.

Statistical Analysis and Machine-Learning Algorithms

All data were analyzed using R version 3.3.3 (R Foundation for Statistical Computing, Vienna, Austria). Kolmogorov−Smirnov normality tests were performed on continuous variables in each group compared to determine normality. Bivariate comparisons were performed using independent sample unpaired Student’s *t* tests for normally distributed continuous variables, and the Wilcoxon rank-sum test was used to compare continuous variables that had a non-normal distribution. Bivariate comparison of categorical variables between groups was done with the χ^2 test. Statistical significance was taken at $p < 0.05$.

The R package Caret was used to build machine-learning models. An initial data partition of both the NIS and the NRD was created, with 70% of the cases allotted to a training dataset and 30% allotted to a testing dataset, with equal event rates in both the training and testing datasets. A repeated cross-validation method was used for the training control, with 10 partitions and 3 repeats, and synthetic minority oversampling technique was used to sample the training data. The synthetic minority oversampling technique was used to mitigate the low event rate of each of our outcomes (25). Two machine-learning classification algorithms were used: logistic regression and gradient boosting. Variable importance is an output for the gradient-boosting algorithm, which is calculated based on how much each variable improves the performance of the algorithm relative to the other variables. Our independent variables for each model were age, elective surgery, sex, insurance status, teaching hospital, deficiency anemias, congestive heart failure (CHF), chronic lung disease, depression, diabetes, hypertension, obesity, peripheral vascular disease (PVD), renal failure, and open fracture. Insurance status was a 4-level factor variable, with private insurance as the reference group. Each machine-learning model was run separately for 3 outcome variables: morbidity and mortality, LOS >3 days, and 30-day readmission. After training, models were then tested on the test datasets. Confusion matrices generated accuracy, sensitivity, and specificity, which were used as performance metrics along with receiver operating characteristic (ROC) curves.

Results

Patient Sample

A total of 50,005 ankle fracture patients who underwent ORIF were analyzed, with 16,501 coming from the NIS and 33,504 from the NRD. Therefore, 16,501 patients were evaluated for developing an inpatient complication or increased LOS, and 33,504 were evaluated for 30-day readmission. A summary of patient and hospital characteristics is provided in Table 1. The rates of morbidity and mortality, LOS >3 days, and 30-day all-cause readmission were 6.9%, 33.7%, and 5.9%, respectively. A breakdown of morbidity and mortality by complication type is presented in Table 2.

Table 1
Summary of demographic, hospital characteristics, and short-term outcomes of the total number of cases isolated from the 2013 and 2014 Nationwide Inpatient Sample and Nationwide Readmissions Database

Characteristic	NIS	NRD
n	16,501	33,504
Age (y)	59.00 (46.00, 71.00)	59.00 (45.00, 71.00)
Elective surgery	2585 (15.7)	5015 (15.0)
Female	11,472 (69.5)	23,156 (69.1)
Insurance		
Private insurance	5384 (32.6)	10,786 (32.2)
Medicaid	1676 (10.2)	3691 (11.0)
Medicare	7037 (42.6)	13,896 (41.5)
Other insurance	2404 (14.6)	5131 (15.3)
Teaching hospital	8958 (54.3)	17,798 (53.1)
Deficiency anemia	1711 (10.4)	3399 (10.1)
Congestive heart failure	971 (5.9)	1711 (5.1)
Chronic lung disease	2707 (16.4)	5280 (15.8)
Depression	2365 (14.3)	4685 (14.0)
Diabetes	2894 (17.5)	5492 (16.4)
Hypertension	8536 (51.7)	16,595 (49.5)
Obesity	2893 (17.5)	5634 (16.8)
Peripheral vascular disease	502 (3.0)	1003 (3.0)
Renal failure	1306 (7.9)	2529 (7.5)
Open fracture	1416 (8.6)	2896 (8.6)
Morbidity and mortality	1136 (6.9)	NA
Length of stay >3 d	5569 (33.7)	NA
30-Day readmission	NA	1969 (5.9)

Abbreviations: NA, not applicable; NIS, Nationwide Inpatient Sample; NRD, Nationwide Readmissions Database.
Data are median (interquartile range) or n (%).

Risk Factors for Adverse Short-Term Outcomes

Several independent risk factors were identified for developing an inpatient complication, increased LOS, or 30-day readmission following ORIF of an ankle fracture (Table 3). Of the 17 independent variables analyzed, older age, nonelective surgery, Medicaid, Medicare, deficiency anemia, CHF, chronic lung disease, diabetes, hypertension, obesity, PVD, and renal failure were found to be statistically significant independent risk factors for developing an inpatient complication. Older age, nonelective surgery, Medicaid, Medicare, other insurance, teaching hospital, deficiency anemia, CHF, chronic lung disease, depression, diabetes, hypertension, obesity, PVD, renal failure, and open fracture were risk factors for increased LOS. Finally, older age, male sex, Medicaid, Medicare, other insurance, nonteaching hospital, deficiency anemia, CHF,

Table 2
Breakdown of morbidity and mortality by complication type (total cases, N = 16,501)

Complication	n (%)
Mechanical wound	24 (0.1)
Postoperative infection	132 (0.8)
Urinary	45 (0.3)
Respiratory	357 (2.2)
Gastrointestinal	52 (0.3)
Cardiovascular	643 (3.9)
Intraoperative	16 (0.1)
Death	29 (0.2)

Table 3
Odds ratios for logistic regression models for each of the 3 outcome variables

	Morbidity and Mortality (n = 16,501)		Length of Stay > 3 Days (n = 16,501)		30-Day Readmission (n = 33,504)	
Characteristic	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
Age (y)	1.025 (1.0194 to 1.0307)	< .0001	1.0273 (1.0248 to 1.0297)	< .0001	1.013 (1.0092 to 1.0168)	< .0001
Elective surgery	0.8328 (0.7003 to 0.9892)	.0378	0.304 (0.2784 to 0.3316)	< .0001	0.9208 (0.809 to 1.0473)	.2102
Female	1.1265 (0.9806 to 1.2945)	.0927	0.9688 (0.9107 to 1.0307)	.3158	0.8764 (0.7936 to 0.9678)	.0092
Medicaid	1.6787 (1.3171 to 2.1363)	< .0001	2.0977 (1.9014 to 2.3143)	< .0001	2.3788 (2.0319 to 2.7852)	< .0001
Medicare	1.6651 (1.4029 to 1.9774)	< .0001	1.5602 (1.4449 to 1.6849)	< .0001	1.9732 (1.7281 to 2.2542)	< .0001
Other insurance	0.9218 (0.7201 to 1.1747)	.5138	1.3342 (1.2164 to 1.4631)	< .0001	1.2614 (1.0741 to 1.4801)	.0045
Teaching hospital	1.12 (0.9916 to 1.2654)	.0684	1.2941 (1.2255 to 1.3667)	< .0001	0.902 (0.8259 to 0.9853)	.022
Deficiency anemia	1.2536 (1.0465 to 1.5018)	.0142	2.0145 (1.8368 to 2.2105)	< .0001	1.406 (1.2299 to 1.6078)	< .0001
Congestive heart failure	1.9623 (1.5864 to 2.4358)	< .0001	2.4553 (2.1549 to 2.8036)	< .0001	1.7734 (1.4952 to 2.1081)	< .0001
Chronic lung disease	1.2598 (1.0768 to 1.4738)	.0039	1.3373 (1.2436 to 1.4382)	< .0001	1.3893 (1.2403 to 1.5563)	< .0001
Depression	1.0785 (0.9122 to 1.2746)	.3758	1.2791 (1.1845 to 1.3812)	< .0001	1.0798 (0.9562 to 1.2192)	0.2151
Diabetes	1.184 (1.0221 to 1.3715)	.0243	1.4268 (1.3281 to 1.533)	< .0001	1.3276 (1.1869 to 1.4851)	< .0001
Hypertension	1.6576 (1.4349 to 1.9153)	< .0001	1.2371 (1.1609 to 1.3183)	< .0001	1.3865 (1.2499 to 1.5379)	< .0001
Obesity	1.5729 (1.3517 to 1.8307)	< .0001	1.6869 (1.5711 to 1.8114)	< .0001	1.0523 (0.9375 to 1.1807)	.3866
Peripheral vascular disease	3.8876 (2.9444 to 5.1962)	< .0001	1.3687 (1.1541 to 1.6268)	3.00E to 04	1.045 (0.8395 to 1.3034)	.6945
Renal failure	1.5022 (1.2305 to 1.8372)	1.0×10^{-4}	1.7497 (1.5668 to 1.9559)	< .0001	1.7735 (1.528 to 2.0608)	< .0001
Open fracture	1.0547 (0.8482 to 1.3104)	.6311	1.1671 (1.061 to 1.2837)	.0015	1.3687 (1.1767 to 1.5921)	< .0001

Abbreviations: CI, confidence interval; OR, odds ratio.

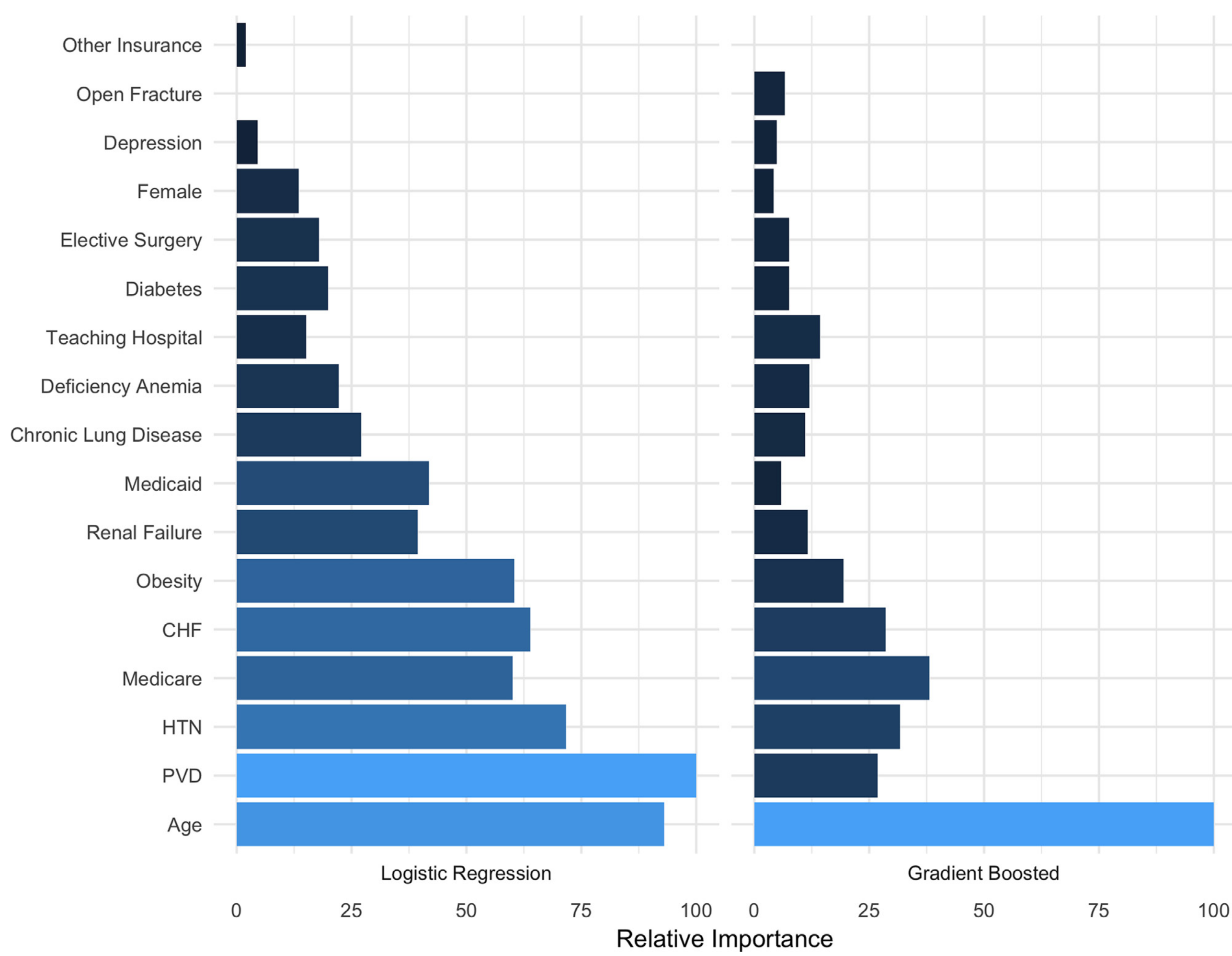

Fig. 1. Variable importance of logistic regression and gradient boosting for predicting morbidity and mortality (n = 16,501). Abbreviations: CHF, congestive heart failure; HTN, hypertension; PVD, peripheral vascular disease.

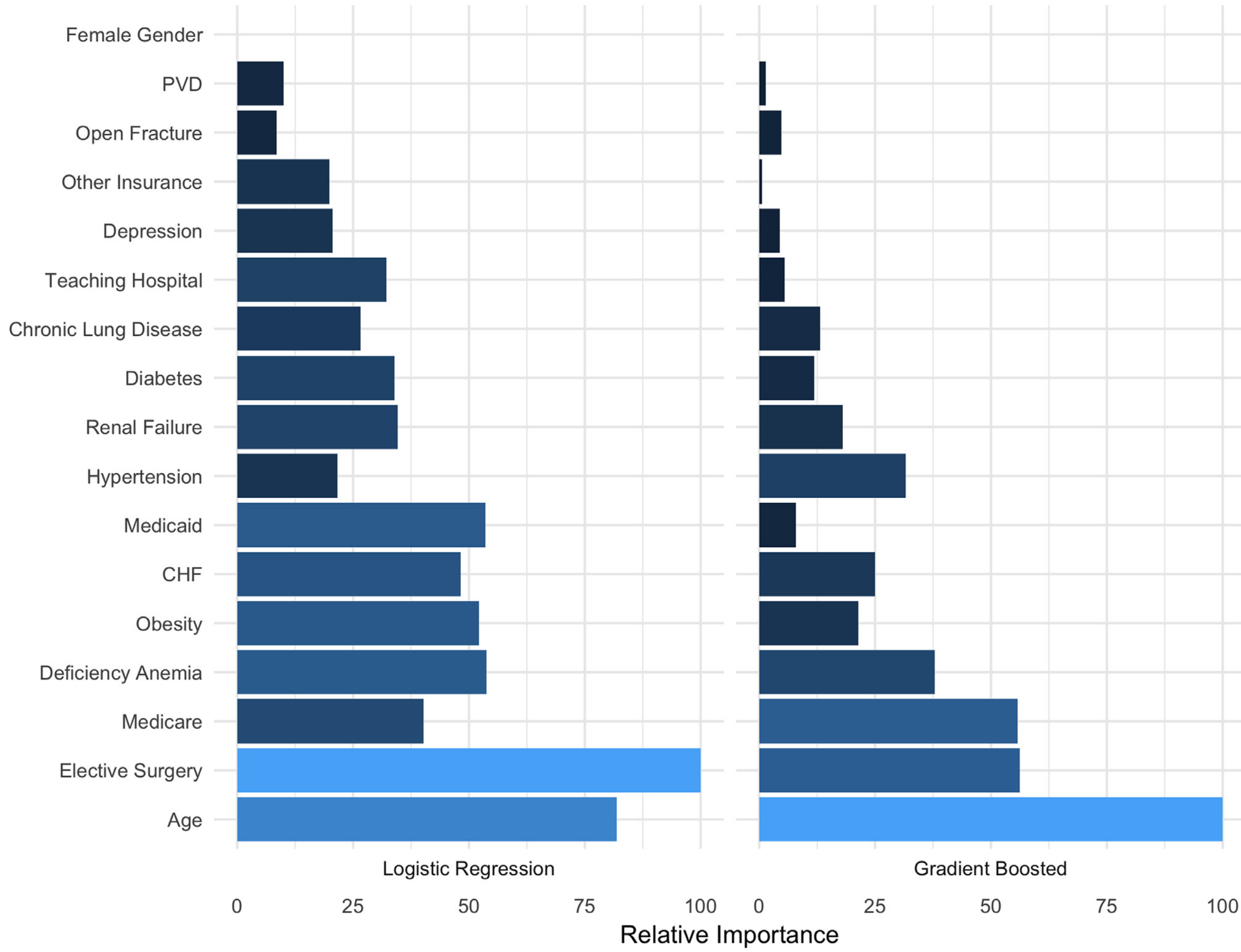

Fig. 2. Variable importance of logistic regression and gradient boosting for predicting length of stay > 3 days (n = 16,501). Abbreviations: CHF, congestive heart failure; PVD, peripheral vascular disease.

chronic lung disease, diabetes, hypertension, renal failure, and open fracture were all risk factors for a 30-day readmission (Table 3).

Variable Importance for Model Prediction

Relative variable importance using the machine-learning algorithms to predict the 3 different outcomes was analyzed. A summary of the relative importance of each variable in the 2 models for predicting morbidity and mortality, LOS > 3 days, and 30-day readmission is presented in Figs. 1, 2, and 3, respectively. Age, hypertension, Medicaid, Medicare, and CHF were consistently among the top variables of importance for prediction of the outcomes by each model.

Comparing Logistic Regression to Gradient Boosting

The gradient-boosting machine-learning model was substantially more accurate than the logistic-regression model for predicting all 3 short-term outcomes, particularly morbidity and mortality and 30-day readmission (Table 4). Additionally, sensitivity was greater for logistic regression for each outcome, while specificity was higher for gradient boosting for each outcome. The ROC area under the curve (AUC) values were comparable between the 2 models, with the logistic regression model having a slightly better value for predicting each outcome (Fig. 4).

Discussion

The objective of the present study was to evaluate the accuracy of machine learning to predict adverse events following ankle fracture ORIF. Several studies have performed risk factor analysis for short-term adverse events following ORIF for ankle fractures and found similar results (8,10,11,26,27). Basques et al (26) conducted a database study using the American College of Surgeons National Quality Improvement Program to identify morbidity and readmissions within 30 days of ORIF for ankle fractures. The authors found that diabetes, older age, American Society for Anesthesiology (ASA) class ≥ 3, bimalleolar fracture, hypertension, and dependent functional status were independent risk factors for a short-term morbid event. Our results corroborate that comorbidities such as diabetes and hypertension may place patients at higher risk of morbidity following ORIF. Basques et al (26) are 1 of the only groups that have investigated 30-day readmission following ORIF of ankle fractures, and they found ASA ≥ 3 to be the only independent risk factor for readmission.

Our results support these findings, because we identified, based on logistic regression, several risk factors for developing morbidity and mortality, having an LOS > 3 days, or experiencing a 30-day readmission. Older age, Medicaid, Medicare, deficiency anemia, CHF, chronic lung disease, diabetes, hypertension, and renal failure were the variables associated with a statistically significant increased risk of

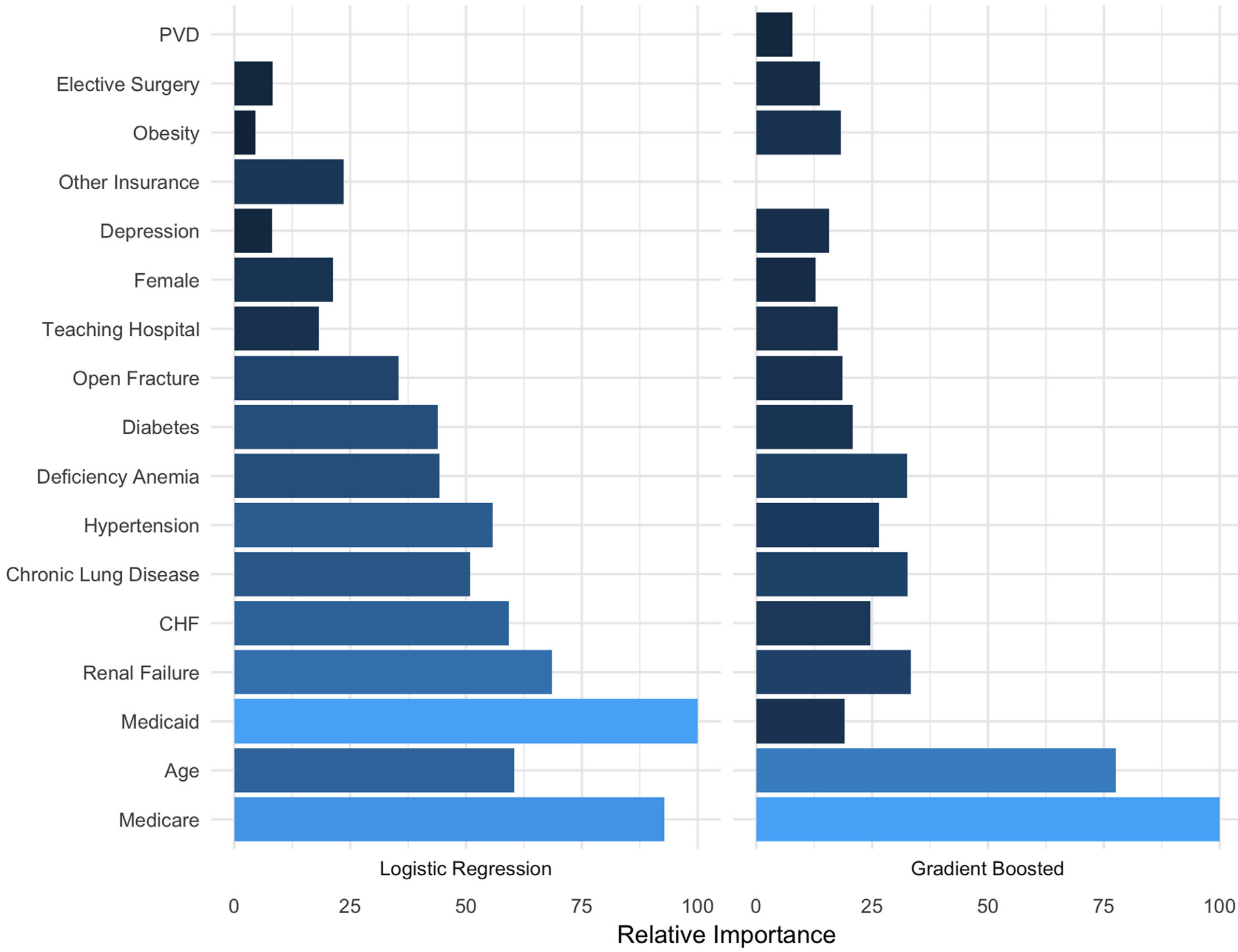

Fig. 3. Variable importance of logistic regression and gradient boosting for predicting 30-day readmission (n = 33,504). Abbreviations: CHF, congestive heart failure; PVD, peripheral vascular disease.

developing all 3 of these postoperative adverse events (Table 3). PVD increased the risk of developing morbidity and mortality by almost 4-fold (odds ratio [OR] 3.89, $p < .0001$) but was not a statistically significant predictor of 30-day readmission (OR 1.04, $p = .69$). Conversely, open fracture was a risk factor for 30-day readmission (OR 1.37, $p < .0001$) but not for morbidity and mortality (OR 1.05, $p = .63$). Female sex (OR 0.88, $p = .01$) and teaching hospital status (OR 0.90, $p = .02$) were associated with a reduced risk of readmission. The literature and our results underscore the importance of evaluating patient comorbidities and counseling patients on risks of adverse short-term outcomes when opting to undergo ORIF for ankle fractures.

Variable importance identifies which variables hold the most relative weight for model prediction. Variable importance provides different information from an output such as an OR, in that it does not provide a direct quantification for the association between a predictor variable and an outcome (27). Rather, variable importance suggests which variables are most important to a predictive model in accurately identifying an outcome. Therefore, the variable importance may identify the most relevant characteristics to consider when evaluating a patient preoperatively. Although the 2 models differed in the variables found to be most important for predicting each of the 3 outcomes, several overlapped between the 2 and between those identified as independent risk factors by logistic regression. Some of these variables included age, insurance status, hypertension, and CHF (Figs. 1–3). These variables were consistently found to be of importance to each of the 2 machine-learning models for predicting all 3 outcomes. These findings, in conjunction with the results of the logistic regression, highlight the impact that age, insurance, and certain comorbidities may have on short-term outcomes after ORIF of ankle fractures. Variable importance has not been reported in the orthopedic literature, but it has been reported in other areas of medicine, suggesting that variable importance of a machine-learning model may provide additional and useful information for identifying characteristics associated with adverse outcomes (27).

Machine learning has become a popular method for building predictive models in the field of medicine and has started to emerge as a risk factor analysis strategy in orthopedics (18–22,27,28). The abundance of

Table 4
Accuracy, sensitivity, specificity, and receiver operating characteristic area under the curve values for logistic regression and gradient boosted machine learning algorithms used to predict 3 different outcomes

	Morbidity and Mortality (n = 16,501)		Length of Stay > 3 Days (n = 16,501)		30-Day Readmission (n = 33,504)	
Algorithm	Logistic	Boosted	Logistic	Boosted	Logistic	Boosted
Accuracy	0.7442	0.8517	0.7149	0.7244	0.7435	0.8479
Sensitivity	0.5753	0.3706	0.5778	0.5497	0.5356	0.3356
Specificity	0.7568	0.8871	0.7847	0.8134	0.7564	0.8798
AUC	0.7529	0.7364	0.7583	0.7580	0.7101	0.6979

Abbreviation: AUC, area under the curve.

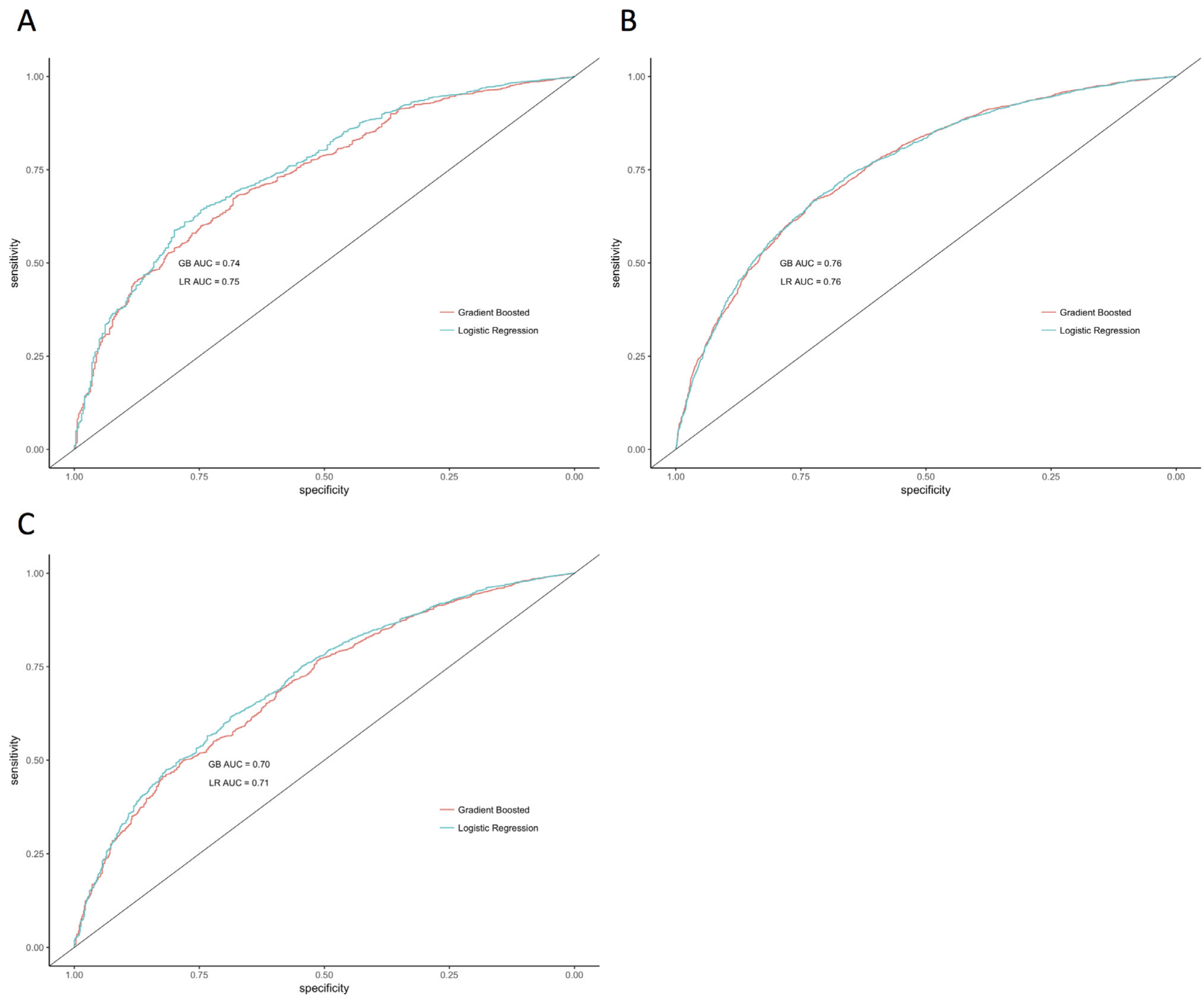

Fig. 4. Receiver operating characteristic curves for each machine learning model predicting (A) morbidity and mortality (n = 16,501), (B) length of stay greater than 3 days (n = 16,501), and (C) 30-day readmission (n = 33,504). Abbreviations: AUC, area under the curve; GB, gradient boosting; LR, logistic regression.

different types of machine learning can make choosing the most accurate algorithm difficult. One of the more commonly used statistical methods for risk factor analysis is logistic regression. Logistic regression is also a type of machine-learning algorithm, and it may therefore serve as a good standard when assessing the performance of other machine-learning algorithms (29,30). Kim et al (22) compared the predictive ability of artificial neural networks to logistic regression for identifying adverse outcomes after posterior lumbar spine surgery. The authors used the 2 algorithms to predict 4 different short-term outcomes: cardiac complications, wound complications, venous thromboembolism, and mortality. They evaluated the performance of the models using sensitivity, specificity, and ROC-AUC. The artificial neural networks had a greater sensitivity than logistic regression for detecting wound complications and mortality, but both models had similar AUC values for all 4 types of complications (Table 4). The authors also compared each model to ASA class as a benchmark predictor and found that each model outperformed ASA with respect to AUC, sensitivity, and specificity. These results suggest that logistic regression and artificial neural networks are comparable machine-learning models, with artificial neural networks providing the benefit of increased sensitivity for certain outcomes (22).

The present study compared a different type of machine learning, gradient boosting, to logistic regression and found similar results to the presented literature. The gradient-boosting models had similar AUC to the logistic-regression models but significantly outperformed the logistic-regression models in terms of overall predictive accuracy and specificity. Overall predictive accuracy is the percentage of correct predictions, and it therefore provides a general picture of the performance of each model. Gradient boosting led to a model that correctly predicted the 3 outcomes more frequently than logistic regression. Gradient boosting may therefore provide the benefit of more accurately identifying patients who will experience an adverse event following ORIF for ankle fractures. However, if a user desires a model that has a higher sensitivity, and may therefore be more tuned to including all patients with adverse events but having more false positives, then an algorithm such as logistic regression may be desirable. The context for which the model will be used must be considered and balanced with the performance statistics of each model. Additional studies in other medical specialties suggest that other machine-learning algorithms perform similarly to, or better than, logistic regression (29,30). Such results, in tandem with our findings, demonstrate the value of exploring different types of machine-learning algorithms when building predictive models.

Machine learning is a powerful tool that can be used by surgeons and patients alike to identify those most at risk of experiencing adverse events following particular types of surgery. Such information would be invaluable for deciding who may be an optimal candidate for surgery, and for when to modify or optimize the surgical care plan for high-risk patients (16). Our findings illustrate that machine learning may be a valid and powerful tool when building predictive models.

This study is limited primarily by the databases used. The NIS and NRD are national databases that collect numerous variables but are not orthopedic focused and do not contain orthopedic-specific variables. As such, important details pertaining to ankle fractures, such as alignment or fracture pattern, were not obtainable. A second limitation of the study is that the 2 databases are not interlinked. Therefore we cannot determine if there is an association between patients who experienced a complication, derived from the NIS, and those who were readmitted, derived from the NRD. Such information could provide valuable insight into the postoperative recovery after ankle fractures. Last, the models developed in the study were validated using these retrospective databases. Future studies should aim to use and compare machine-learning models on prospective cohorts.

In conclusion, several patient characteristics and comorbidities were identified as risk factors for developing morbidity and mortality, LOS > 3 days, or 30-day readmission. Some of the most consistent risk factors included age, insurance status, hypertension, and CHF, among many others. In comparing the performance of 2 machine-learning algorithms for predicting outcomes, gradient boosting was found to be more accurate and more specific than logistic regression, but the models had similar ROC-AUC values. These results demonstrate the importance of evaluating patients preoperatively for factors that may put them at higher risk of developing an adverse event and suggest that machine learning can accurately predict short-term outcomes after ORIF of ankle fractures.

References

1. Salai M, Dudkiewicz I, Novikov I, Amit Y, Chechick A. The epidemic of ankle fractures in the elderly—is surgical treatment warranted? Arch Orthop Trauma Surg 2000;120:511–513.
2. Shibuya N, Davis ML, Jupiter DC. Epidemiology of foot and ankle fractures in the United States: an analysis of the National Trauma Data Bank (2007 to 2011). J Foot Ankle Surg 2014;53:606–608.
3. Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. Injury 2006; 37:691–697.
4. Court-Brown CM, McBirnie J, Wilson G. Adult ankle fractures—an increasing problem? Acta Orthop Scand 1998;69:43–47.
5. Van Schie-Van der Weert EM, Van Lieshout EMM, De Vries MR, Van der Elst M, Schepers T. Determinants of outcome in operatively and non-operatively treated Weber-B ankle fractures. Arch Orthop Trauma Surg 2012;132:257–263.
6. Pakarinen HJ, Flinkkil TE, Ohtonen PP, Ristiniemi JY. Stability criteria for nonoperative ankle fracture management. Foot Ankle Int 2011;32:141–147.
7. Yang E, Wu Y, Dorcil J. Surgical versus nonsurgical treatment of the SE4-equivalent ankle fracture: a retrospective functional outcome study. Orthopedics 2011;34:271.
8. Cavo MJ, Fox JP, Markert R, Laughlin RT. Association between diabetes, obesity, and short-term outcomes among patients surgically treated for ankle fracture. J Bone Joint Surg Am 2015;97:987–994.
9. Leyes M, Torres R, Guillén P. Complications of open reduction and internal fixation of ankle fractures. Foot Ankle Clin 2003;8:131–147, ix.
10. Miller AG, Margules A, Raikin SM. Risk factors for wound complications after ankle fracture surgery. J Bone Joint Surg Am 2012;94:2047–2052.
11. SooHoo NF, Krenek L, Eagan MJ, Gurbani B, Ko CY, Zingmond DS. Complication rates following open reduction and internal fixation of ankle fractures. J Bone Joint Surg Am 2009;91:1042–1049.
12. Belmont PJ, Davey S, Rensing N, Bader JO, Waterman BR, Orr JD. Patient-based and surgical risk factors for 30-day postoperative complications and mortality after ankle fracture fixation. J Orthop Trauma 2015;29:e476–e482.
13. Brasier AR, Ju H. Analysis and predictive modeling of asthma phenotypes. Adv Exp Med Biol 2014;795:273–288.
14. Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. Endocrinol Metab (Seoul) 2016;31:38–44.
15. Yankeelov TE, An G, Saut O, Luebeck EG, Popel AS, Ribba B, Vicini P, Zhou X, Weis JA, Ye K, Genin GM. Multi-scale modeling in clinical oncology: opportunities and barriers to success. Ann Biomed Eng 2016;44:2626–2641.
16. Osorio JA, Scheer JK, Ames CP. Predictive modeling of complications. Curr Rev Musculoskelet Med 2016;9:333–337.
17. Baştanlar Y, Ozuysal M. Introduction to machine learning. Methods Mol Biol 2014; 1107:105–128.
18. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: a literature review. Front Bioeng Biotechnol 2018;6:75.
19. Pedoia V, Haefeli J, Morioka K, Teng HL, Nardo L, Souza RB, Ferguson AR, Majumdar S. MRI and biomechanics multidimensional data analysis reveals R2-R1p as an early predictor of cartilage lesion progression in knee osteoarthritis. J Magn Reson Imaging 2018;47:78–90.
20. Dolatabadi E, Taati B, Mihailidis A. An automated classification of pathological gait using unobtrusive sensing technology. IEEE Trans Neural Syst Rehabil Eng 2017;25: 2336–2346.
21. Matic A, Petrovic Savic S, Ristic B, Stevanovic CB, Devedzic G. Infrared assessment of knee instability in ACL deficient patients. Int Orthop 2016;40:385–391.
22. Kim JS, Merrill RK, Arvind V, Kaji D, Pasik SD, Nwachukwu CC, Vargas L, Osman NS, Oermann EK, Caridi JM, Cho SK. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. Spine 2017;43:853–860.
23. HCUP National Inpatient Sample. Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD; 2012.
24. HCUP Nationwide Readmissions Database (NRD). Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD; 2012.
25. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority oversampling technique. J Artif Intell 2012;16:321–357.
26. Basques BA, Miller CP, Golinvaux NS, Bohl DD, Grauer JN. Morbidity and readmission after open reduction and internal fixation of ankle fractures are associated with preoperative patient characteristics. Clin Orthop 2015;473:1133–1139.
27. Koval KJ, Zhou W, Sparks MJ, Cantu RV, Hecht P, Lurie J. Complications after ankle fracture in elderly patients. Foot Ankle Int 2007;28:1249–1255.
28. Lazzarini N, Runhaar J, Bay-Jensen AC, Thudium CS, Bierma-Zeinstra SMA, Henrotin Y, Bacardit J. A machine learning approach for the identification of new biomarkers for knee osteoarthritis development in overweight and obese women. Osteoarthritis Cartilage 2017;25:2014–2021.
29. Stylianou N, Akbarov A, Kontopantelis E, Buchan I, Dunn KW. Mortality risk prediction in burn injury: comparison of logistic regression with machine learning approaches. Burns J Int Soc Burn Inj 2015;41:925–934.
30. Henrard S, Speybroeck N, Hermans C. Classification and regression tree analysis vs. multivariable linear and logistic regression methods as statistical tools for studying haemophilia. Haemoph Off J World Fed Hemoph 2015;21:715–722.