Examining the Ability of Artificial Neural Networks Machine Learning Models to Accurately Predict Complications Following Posterior Lumbar Spine Fusion

Jun S. Kim, MD[*], **Robert K. Merrill, BS**[*], **Varun Arvind, BS**[*], **Deepak Kaji, BA**[*], **Sara D. Pasik, BA**[*], **Chuma C. Nwachukwu, BA**[*], **Luilly Vargas, BSN**[*], **Nebiyu S. Osman, BA**[*], **Eric K. Oermann, MD**[†], **John M. Caridi, MD**[†], and **Samuel K. Cho, MD**[*]

[*]Department of Orthopedic Surgery, Icahn School of Medicine at Mount Sinai, New York, NY

[†]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York, NY

Abstract

Study Design.—A cross-sectional database study.

Objective.—The aim of this study was to train and validate machine learning models to identify risk factors for complications following posterior lumbar spine fusion.

Summary of Background Data.—Machine learning models such as artificial neural networks (ANNs) are valuable tools for analyzing and interpreting large and complex datasets. ANNs have yet to be used for risk factor analysis in orthopedic surgery.

Methods.—The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) database was queried for patients who underwent posterior lumbar spine fusion. This query returned 22,629 patients, 70% of whom were used to train our models, and 30% were used to evaluate the models. The predictive variables used included sex, age, ethnicity, diabetes, smoking, steroid use, coagulopathy, functional status, American Society for Anesthesiology (ASA) class 3, body mass index (BMI), pulmonary comorbidities, and cardiac comorbidities. The models were used to predict cardiac complications, wound complications, venous thromboembolism (VTE), and mortality. Using ASA class as a benchmark for prediction, area under receiver operating curves (AUC) was used to determine the accuracy of our machine learning models.

Results.—On the basis of AUC values, ANN and LR both outperformed ASA class for predicting all four types of complications. ANN was the most accurate for predicting cardiac complications, and LR was most accurate for predicting wound complications, VTE, and mortality, though ANN and LR had comparable AUC values for predicting all types of complications. ANN had greater sensitivity than LR for detecting wound complications and mortality.

Conclusion.—Machine learning in the form of logistic regression and ANNs were more accurate than benchmark ASA scores for identifying risk factors of developing complications following posterior lumbar spine fusion, suggesting they are potentially great tools for risk factor analysis in spine surgery.

Keywords

AI; ANN; artificial intelligence; artificial neural networks; complications; logistic regression; LR; machine learning; neural nets; PLF; posterior lumbar fusion; prediction

---

Posterior lumbar spine fusion is a surgical strategy used to treat various degenerative conditions of the lumbar spine.[1] Posterior lumbar fusion is known to be an effective procedure, but is associated with several complications that can lead to adverse outcomes.[2–6] Risk factors for these adverse outcomes can be identified using machine learning.

Machine learning has been used in a wide variety of applications, ranging from credit card fraud detection to online advertising. It provides a computational system that uses data from which to continually learn, develop algorithms, and make predictions. Machine learning classifiers can perform these functions without prior assumptions, leading to a highly adaptable system with minimal bias.[7] Clinical medicine requires physicians to handle enormous quantities of complex data. Machine learning provides the opportunity to analyze such data, and has an advantage over human-based computations because machine learning algorithms have a greater capability of identifying unintuitive patterns in large patient datasets.[8]

Currently, multivariable logistic regression (LR), a form of machine learning, is one of the most commonly used methods for identifying risk factors predictive of developing complications.[9–12] Artificial neural networks (ANNs) are another type of machine learning that, in contrast to LR, are nonlinear and more flexible, which may allow for the identification of nonlinear patterns that make predictions more accurate.[13] ANNs are beginning to take hold in the field of medicine, and could provide valuable insight into current evidence-based medicine.[14,15] In the burgeoning era of rising health care costs and greater scrutiny over surgical outcomes, there has been an increasing emphasis on understanding the risk factors and possible predictors to optimize perioperative planning and management. Data-driven clinical decision support tools have the potential to lead to cost savings by leveraging the information contained in large medical databases. Uptake of machine learning approaches in this realm has lagged due to unfamiliarity and sparse data sets.

This study seeks to develop and validate ANN machine learning algorithms to precisely predict complications following posterior lumbar fusion using a national database. These algorithms have the capability of continuously "learning" using newly generated information to improve the quality and efficiency of care. We hypothesized that machine learning techniques applied in other fields may be equally capable for use in predicting postoperative complications following posterior lumbar fusion.

MATERIALS AND METHODS

Patient Selection

The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) database was used to train and validate ANN and LR models. ACS-NSQIP is a prospective, risk-adjusted, multicenter quality improvement program that prospectively collects more than 135 preoperative, intraoperative, and 30-day postoperative outcomes from operative reports, medical records, and patient interviews from patients undergoing surgical procedures in both the in-patient and out-patient setting. The data are collected from over 258 participating hospitals within the United States with an interrater reliability of 1.56% as of 2008 as has been reported.[16] More information can be found at http://www.acsnsqip.org. The database was queried for patients who underwent posterior lumbar spine fusion between 2010 and 2014, and patients with missing data were excluded from the study.

Training and Hold-Out Data Sets

We used 70% of the initial data for training our models and 30% for post-training evaluation of the models (Figure 1A, B). Adaptive synthetic sampling approach to imbalanced learning (ADASYN) was used to generate positive complications as a means to overcome the low incidence of complications. Adaptive synthetic sampling uses a weighted distribution for minority class examples that are difficult to learn, generating synthetic data based on such examples to improve the mode learning and generalizability.[17]

Feature Selection

Input features used for training include sex, age, ethnicity (white, black, Hispanic, or other), history of diabetes, history of smoking, steroid use, history of bleeding disorders, function status, American Society for Anesthesiology (ASA) class 3, body mass index (BMI), pulmonary comorbidities, and cardiac comorbidities. The number of training examples required to reach a certain accuracy grows exponentially with the number of irrelevant features; therefore, we performed feature selection as a means to improve the generalizability of our ANN and LR models. Stepwise LR was performed using a MatLab in-built function on the training data set to obtain probability coefficients for each feature. The top six features with the greatest regression coefficient magnitudes, with $P < 0.05$ were used as input variables for the ANN and LR for mortality, venous thromboembolism (VTE), cardiac complications, and wound complications, respectively. Age and BMI were considered continuous variables, while all others were considered categorical.

Machine Learning Construction and Testing

Machine learning models were trained to predict occurrence of mortality, VTE, wound complications, and cardiac complications encoded as binary outcome variables. ANNs were designed using the Neural Network toolbox in MatLab 2016b (MathWorks, Inc., Natick, MA). L2 regularization was used to avoid ANN overfitting, by augmenting the error function used for training with the squared magnitude of the weights used in the ANN. This ensures that overly complex models are not overfitted to a specific dataset, thus improving predictive generalizability. Due to the large class imbalance even post-ADASYN, multiple

ANNs were created by partitioning the majority class into sub sets in a 1:1 ratio with the minority class, creating ANNs trained off each partition. Patients were subsequently distributed in a 2:1:1 ratio to generate training, validation, and testing data sets, respectively. Hold-out data not used for training or testing were used for final testing of the ANN to provide an unbiased assessment of ANN performance. Hold-out data were input into each ANN, and final predictions were based on individual accuracy-weighted predictions surveyed across each ANN. The dataset was randomized and repartitioned and training and testing was repeated five times to evaluate statistical significance of ANN performance.

Statistical Analysis

ANN performance was compared to LR that was trained and tested on the same data by which the ANN was evaluated. These two machine learning models were also compared with ASA scoring based on area under the receiver operating characteristic curve (AUC), which is the gold-standard metric for evaluating machine learning algorithms with 95% confidence intervals (95% CIs). A value of 0.50 for AUC indicates random chance for identifying a complication.

RESULTS

Cases and Complications

We analyzed a total of 22,629 patients following our exclusion criteria with an average age of 60 years. Table 1 summarizes the patient characteristics and number of complications that occurred. The most common complications were wound related, occurring in 2.2% of cases. For cardiac, VTE, wound complications, and mortality training sets 937, 556, 884, and 1009 cases were excluded, respectively, due to incomplete medical records. Adaptive synthetic sampling generated 779 cases for the cardiac training set, 749 cases for the VTE set, 659 cases for the wound complications set, and 799 for the mortality set. Age, diabetes, and BMI were found to be strong predictors of developing a complication, particularly for mortality (Figure 2).

LR, ANN, and ASA Performance

On the basis of the AUC with ASA as the benchmark, LR and ANNs were found to outperform ASA for predicting each type of complication (Figures 3, 4). ANN had the best AUC for predicting cardiac complications, while LR had the best AUC value for predicting VTE, wound complications, and mortality sets (Table 2). With respect to mortality and wound complications, ANN was more sensitive than LR for predicting such complications (Figure 5A, B).

DISCUSSION

Risk factor analysis has become an important aspect of medical research for a variety of fields. Large sample studies with high quality data offer the opportunity to identify patients and surgical factors that may increase the risk of developing a surgical complication. The objective of this study was to take risk factor analysis a step further by determining whether

ANNs could accurately predict if posterior lumbar fusion patients would develop a surgical complication based on patient features.

Previous groups have employed the use of ANNs and other machine learning models to these data sets. [18–21] However, these studies either trained models on extremely large databases (>1,400,000 patients) or on complications with high occurrence rates. These examples are impractical for independent institutions or for small scale procedures with rare complications. Low occurrence rates in relatively small datasets lead to large class-imbalances that are a significant challenge in medical machine learning.[22,23] To this end, we have trained several supervised machine learning classifiers to predict the probability of postoperative complications in a relatively small dataset (<15,000 patients) that can accurately learn complications with relatively low occurrence rates (<1%). We have rigorously developed and tested our models by employing the best practices in machine learning in this study by performing automated feature selection, L2 regularization, testing on blinded hold-out data sets, and comparing to a standard risk-scoring system to ensure a high standard that is necessary for implementation of machine learning in clinical settings.

ANNs differ from LR models in that they are nonlinear. This allows for more flexible modeling, which can lead to the identification of complex patterns.[13] Dreiseitl and Ohno-Machado[13] reviewed these two machine learning models and identified 72 studies that had compared the two based on the ability to discriminate two datasets. The authors found that when statistical testing was performed to compare the two models, ANN was superior 18% of the time, LR 1% of the time, and no difference 42% of the time. When no statistical testing was performed, ANNs were better 33% of the time, LR was better 6% of the time, and not difference 0% of the time. These results suggest that the two forms of machine learning may be comparable, but that ANNs may have a slightly better accuracy.

Our results corroborate these findings. ASA has been shown to be a risk factor for developing complications after lumbar spine surgery and we therefore chose it as a benchmark.[7] Our AUC values demonstrate the accuracy of ASA class, ANN, and LR for predicting cardiac complications, VTE, wound complications, and mortality. Both ANN and LR outperformed ASA for predicting all types of complications. ANN outperformed LR for cardiac complication prediction and LR outperformed ANN for VTE, wound complications, and mortality, though both models had relatively similar AUC values. These results suggest that the two methods are comparable, though for certain complications one may be slightly superior.

When examining our confusion matrices, we found ANN to have greater sensitivity than LR for detecting mortality and wound complications, and LR to have greater specificity for mortality and wound complications. Part of the reason we may have observed this is that the ANN was better able to handle the large class-imbalance inherent in our data set and was able to better generalize its predictions when evaluated on the blinded hold-out data set. L2 regularization was performed on the ANN to prevent overfitting, and this may explain the improved sensitivity of ANN when compared with LR. In addition, ANNs have reported better ability to predict nonlinear patterns contained within the dataset compared with LR.[24] This suggests that ANNs may be a more suitable machine learning platform for use in the

clinical setting, as it can be rigorously developed to ensure a high degree of sensitivity that is important for medical prognostication.

To summarize, both LR and ANN are valid methods of predicting complications, and seem to be comparable to each other. ANN may provide the advantage of having an increased sensitivity for certain adverse events such as mortality and wound complications. Furthermore, we demonstrate that machine learning approaches can be applied to various health systems with or without organized databases using approaches outlined in this study. This expands the use of such predictive algorithms to diverse health care settings. The ability of machine learning to identify at-risk patients and predict potential complications has been clearly demonstrated here, yet the ability to suggest avenues of treatment based on predicted complications has not yet been realized. Future work can take advantage of electronic medical registries and medical literature to suggest optimal treatment strategies based on key patient data. Such models can not only guide physicians in the decision making process but can also aide health care systems in low-resource settings, provide personalized care, and improve response times during critical settings. Taken together, the opportunities described here can be used to strengthen medical artificial intelligence (AI) to improve surgical outcomes.

Some limitations of this study should be noted, the majority of which are related to weaknesses within the NSQIP database we chose to use. The performance of any classifier is rooted, in part, in the quality of the training data. Therefore, weaknesses in the NSQIP are represented as weaknesses in the neural network classifier. Larger national in-patient datasets such as the National In-Patient Sample (NIS) exist. Such data sets sampling patients with a broad demographic spectrum can serve to elucidate patterns in the model that are both more generalizable and predictive of future complications and risk. A major challenge in medicine is the paucity of highly granular and robust large-scale datasets for specific operational cohorts. Large-scale databases remain scattered across institutions and are isolated to protect patient privacy.[25] Furthermore, the NSQIP dataset was not designed with spine surgery outcomes in mind. As a result, many features or patient variables, which may serve as stronger inputs, were not available. Further, this study does not differentiate between different approaches such as minimally invasive *versus* open lumbar fusion that are significant modifiers for type and probability of postoperative complications. Particularly, minimally invasive posterior lumbar fusion was found to have significantly less blood loss, recovery time, and length of hospital stay all three of which are important risk contributors to post-operative complications.[26] Future more advanced models should aim at incorporating such distinctions.

CONCLUSION

ANNs are a valid method for identifying risk factors of developing complications following posterior lumbar spine fusion. Both LR and ANN were more accurate than benchmark ASA scores based on AUC values. In the past, generalized linear models such as the LR have been the most commonly used classifiers for this purpose. However, the machine learning models described here, particularly the ANN, are similarly powerful and in some circumstances, far exceed LR. As the ability to obtain high-quality patient data and

computing power increases over time, it is likely that machine learning techniques will find themselves increasingly commonplace in the hospital setting.

Acknowledgments

References

1. Mobbs RJ, Phan K, Malham G, et al. Lumbar interbody fusion: techniques, indications and comparison of interbody fusion options including PLIF, TLIF, MI-TLIF, OLIF/ATP, LLIF and ALIF. J Spine Surg 2015;1:2–18. [PubMed: 27683674]

2. Rompe JD, Eysel P, Hopf C. Clinical efficacy of pedicle instrumentation and posterolateral fusion in the symptomatic degenerative lumbar spine. Eur Spine J 1995;4:231–7. [PubMed: 8528782]

3. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. Eur Spine J 2001;10:309–13. [PubMed: 11563616]

4. White SF, Asher MA, Lai SM, et al. Patients' perceptions of overall function, pain, and appearance after primary posterior instrumentation and fusion for idiopathic scoliosis. Spine (Phila Pa 1976) 1996;24:1699–700.

5. Lowe TG, Tahernia AD, O'Brien MF, et al. Unilateral transforaminal posterior lumbar interbody fusion (TLIF): indications, technique, and 2-year results. J Spinal Disord Tech 2002;15:31–8. [PubMed: 11891448]

6. Freeman BJC, Licina P, Mehdian SH. Posterior lumbar interbody fusion combined with instrumented postero-lateral fusion: 5-year results in 60 patients. Eur Spine J 2000;9:42–6. [PubMed: 10766076]

7. Shavlik J, Dietterich T, editors. Readings in Machine Learning. San Mateo, CA: Morgan Kaufmann; 1990.

8. Obermeyer Z, Emanuel EJ. Predicting the future: big data, machine learning, and clinical medicine. N Engl J Med 2016;375:1216–9. [PubMed: 27682033]

9. Capua JD, Somani S, Kim JS, et al. Analysis of risk factors for major complications following elective posterior lumbar fusion. Spine (Phila Pa 1976) 2017;42:1347–54. [PubMed: 28146019]

10. Lee NJ, Kothari P, Phan K, et al. The incidence and risk factors for 30-day unplanned readmissions after elective posterior lumbar fusion. Spine (Phila Pa 1976) 2018;43:41–8. [PubMed: 27031773]

11. Pugely AJ, Martin CT, Gao Y, et al. Causes and risk factors for 30-day unplanned readmissions after lumbar spine surgery. Spine (Phila Pa 1976) 2014;39:761–8. [PubMed: 24525993]

12. Akins PT, Harris J, Alvarez JL, et al. Risk factors associated with 30-day readmissions after instrumented spine surgery in 14,939 patients: 30-day readmissions after instrumented spine surgery. Spine (Phila Pa 1976) 2015;40:1022–32. [PubMed: 25856262]

13. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002;35:352–9. [PubMed: 12968784]

14. Kononenko I Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med 2001;23:89–109. [PubMed: 11470218]

15. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2006;2:59–77.

16. Ingraham AM, Richards KE, Hall BL, et al. Quality improvement in surgery: the American College of Surgeons national surgical quality improvement program approach. Adv Surg 2010;44:251–67. [PubMed: 20919525]

17. He H, Bai Y, Garcia EA, et al. ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence); June 8, 2008 1322–1328.

18. Van Esbroeck A, Rubinfeld I, Hall B, et al. Quantifying surgical complexity with machine learning: looking beyond patient factors to improve surgical models. Surgery 2014;156:1097–105. [PubMed: 25108343]

19. Hu Z, Simon GJ, Arsoniadis EG, et al. Automated detection of postoperative surgical site infections using supervised methods with electronic health record data. Stud Health Technol Inform 2015;216:706–10. [PubMed: 26262143]

20. Sohn S, Larson DW, Habermann EB, et al. Detection of clinically important colorectal surgical site infection using Bayesian network. J Surg Res 2017;209:168–73. [PubMed: 28032554]

21. Bilimoria KY, Liu Y, Paruch JL, 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;217:833–842.e1–e3. [PubMed: 24055383]

22. Krell MM, Wilshusen N, Seeland A, et al. Classifier transfer with data selection strategies for online support vector machine classification with class imbalance. J Neural Eng 2017;14:025003. [PubMed: 28192282]

23. Wang Q, Luo Z, Huang J, et al. A novel ensemble method for imbalanced data learning: bagging of extrapolation-SMOTE SVM. Comput Intell Neurosci 2017;2017:1827016. [PubMed: 28250765]

24. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49:1225–31. [PubMed: 8892489]

25. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. JAMA 2014;311:2479–80. [PubMed: 24854141]

26. Park Y, Ha JW. Comparison of one-level posterior lumbar interbody fusion performed with a minimally invasive approach or a traditional open approach. Spine (Phila Pa 1976) 2007;32: 537–43. [PubMed: 17334287]

Key Points

□ We conducted a cross-sectional database study to train and validate two types of machine learning models, logistic regression and artificial neural networks, for predicting complications following posterior lumbar spine fusion.

□ We analyzed a total of 22,629 patients, and used American Society of Anesthesiologists (ASA) class as a benchmark for predicting cardiac complications, wound complications, venous thromboembolism, and mortality.

□ We found that artificial neural networks outperformed ASA class and sometimes logistic regression for predicting these four types of complications.
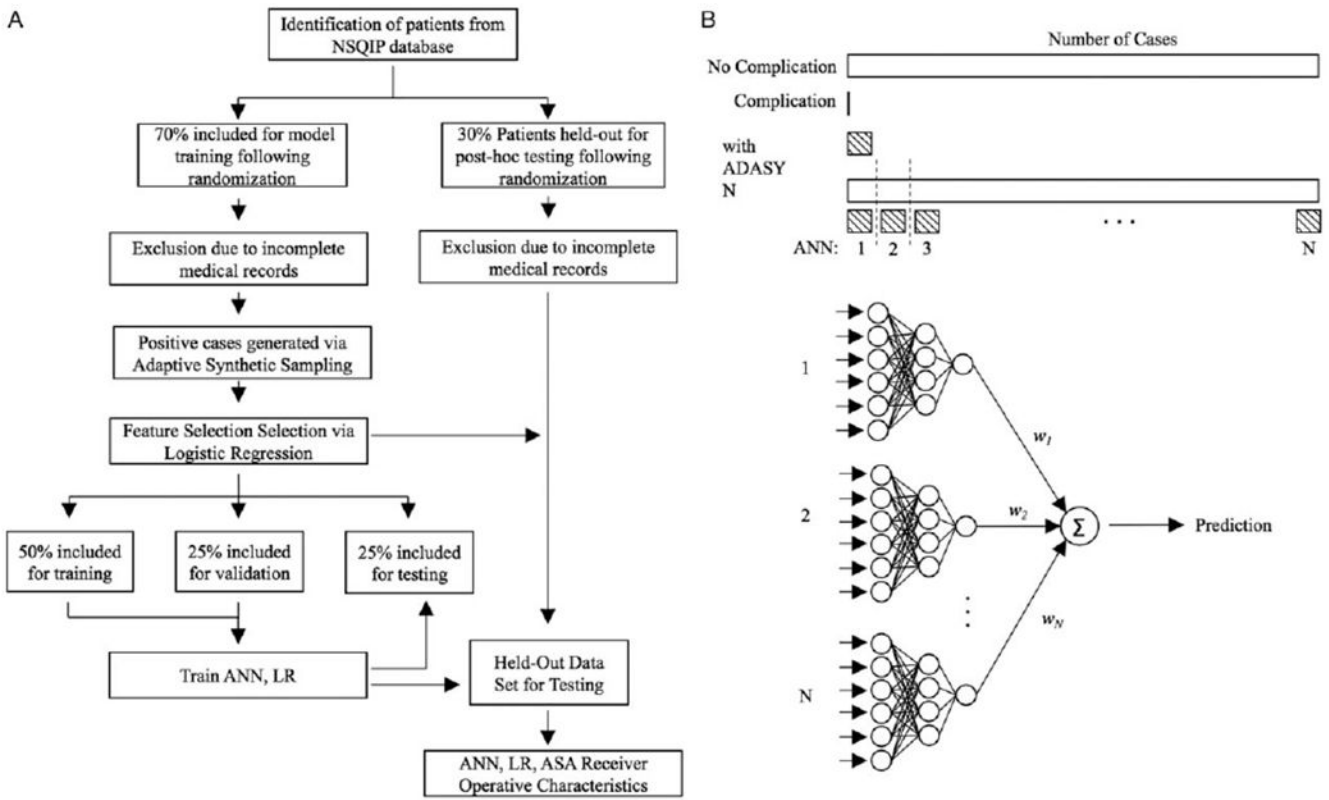
Figure 1.
(**A**) Schematic of study workflow. (**B**) Diagram of ANN model. Bar lengths represent number of patient cases. ADASYN increases the number of positive cases to combat class imbalance. Negative cases are then partitioned in a 1:1 ratio with the positive cases to create a class-balanced dataset used for ANN training. Each partition trains an independent neural net. During evaluation, data are fed through each neural net where the responses are surveyed, weighted by the model's accuracy, and the net prediction is used.
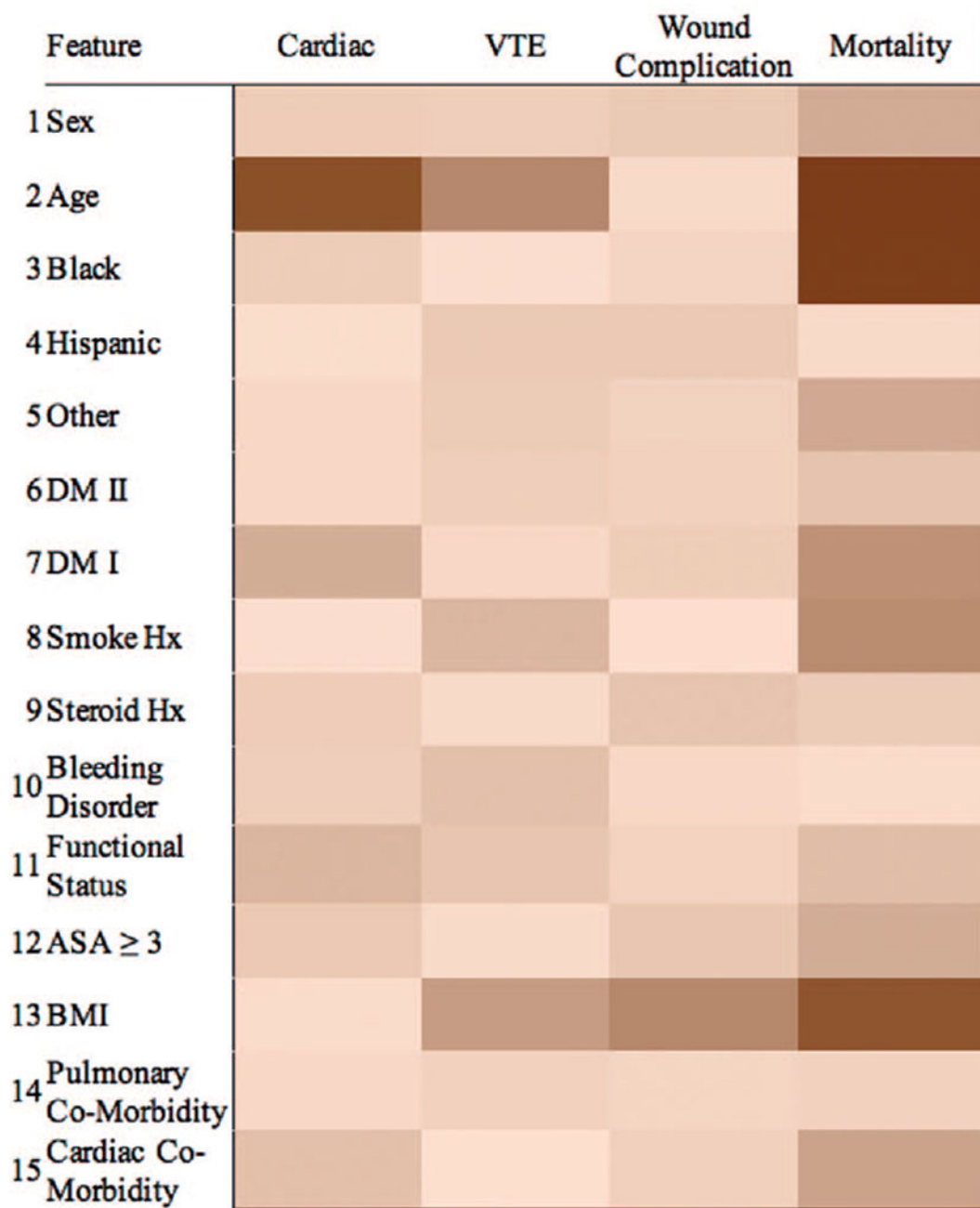
Figure 2.
Coefficient weights obtained from logistic regression analysis used for feature selection.
Dark cells indicate highly weighted features indicating a strong predictive value, and lighter
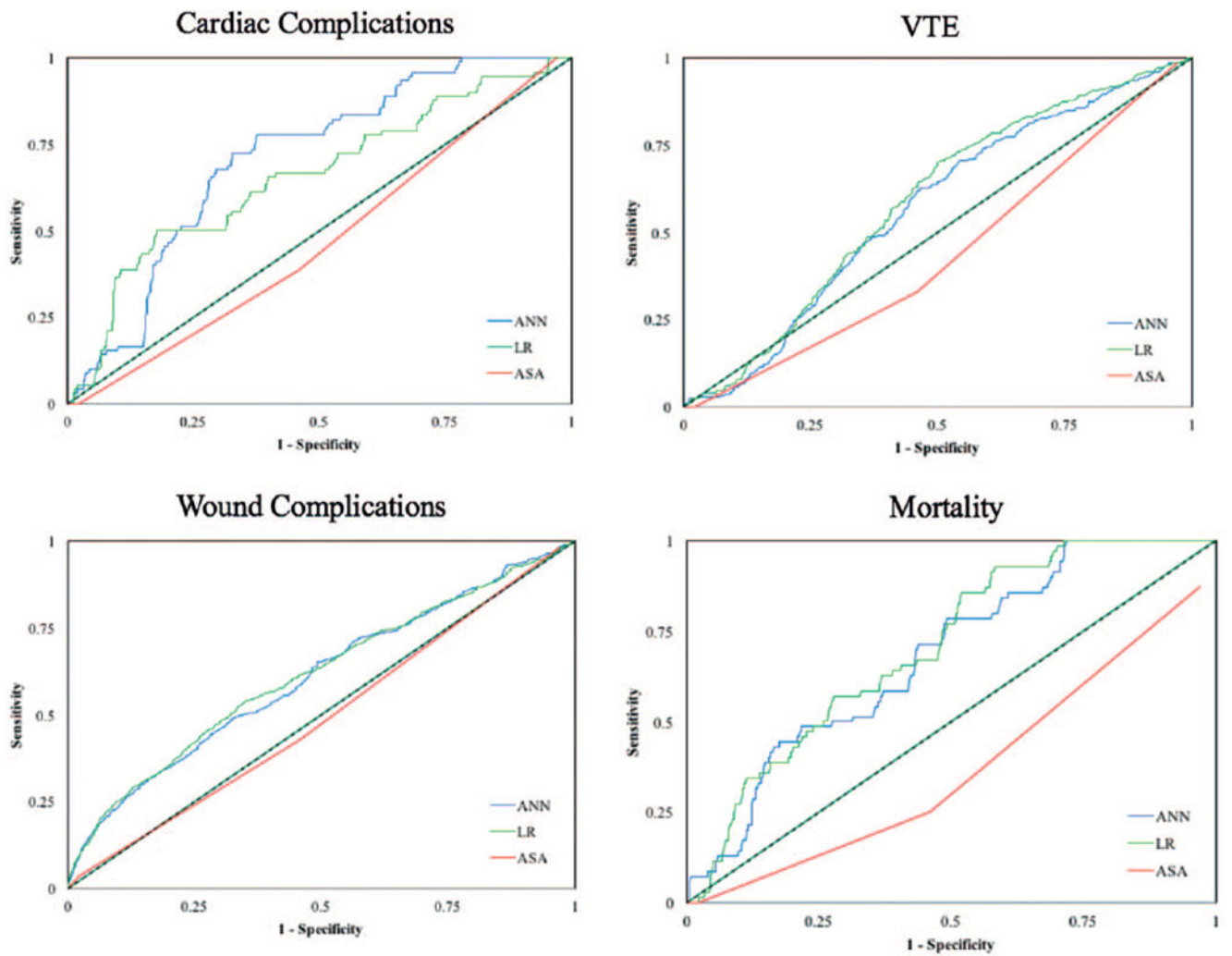cells indicate weakly weighted features.

Figure 3.
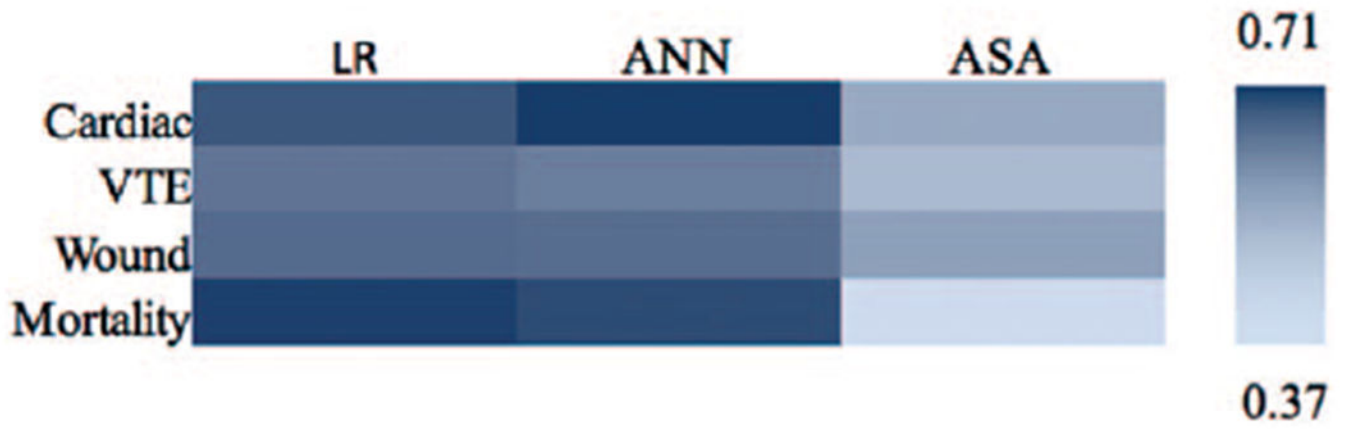Receiver operating curves for ASA, LR, and ANN for each complication type.

Figure 4.
Heatmap of AUC values from LR, ANN, and ASA when predicting cardiac complications
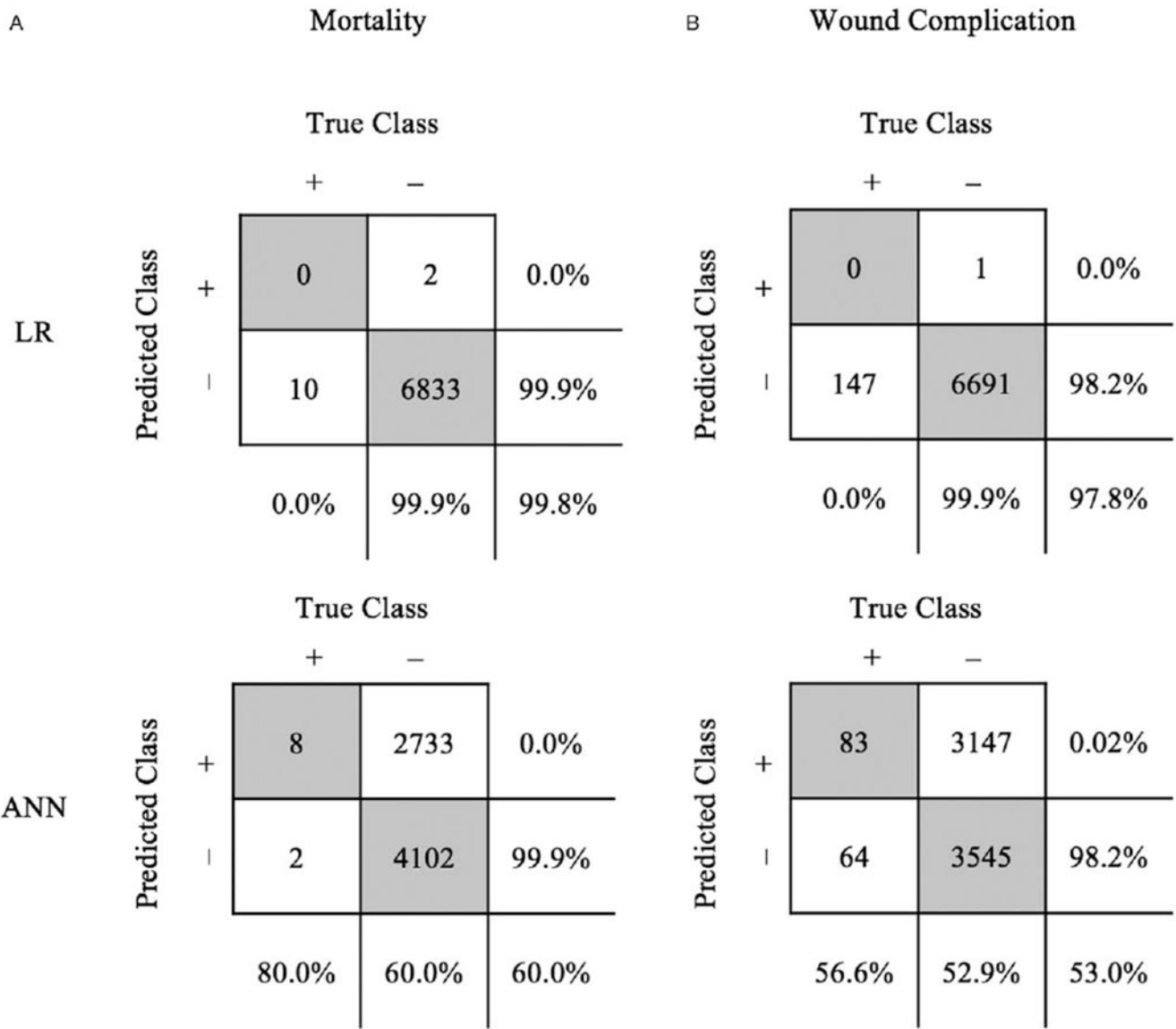(cardiac), VTE, wound complications (wound), and mortality.

Figure 5.
Confusion matrices of trained ANN and LR machine learners evaluated on hold-out (**A**) mortality and (**B**) wound complication data sets to demonstrate real-world performance.

TABLE 1.

Summary of Patient Characteristics and Complications

Feature	Average	Total	Cardiac Complication	VTE Complication	Wound Complication	Mortality
Sex, n (%)						
Male		10191 (45.0)	47 (0.5)	109 (1.1)	191 (1.9)	20 (0.2)
Female		12438 (55.0)	53 (0.4)	113 (0.9)	299 (2.4)	14 (0.1)
Age, mean	60.2					
Ethnicity, n (%)						
White		19197 (84.8)	87 (0.5)	190 (1.0)	396 (2.1)	31 (0.2)
Black		1605 (7.1)	4 (0.2)	17 (1.1)	47 (2.9)	0 (0.0)
Hispanic		196 (0.9)	2 (1.0)	2 (1.0)	2 (1.0)	1 (0.5)
Other		1631 (7.2)	7 (0.4)	13 (0.8)	45 (2.8)	2 (0.1)
Diabetes mellitus, n (%)						
No		18679 (82.5)	71 (0.4)	181 (1.0)	374 (2.0)	27 (0.1)
Type II		2703 (11.9)	13 (0.5)	25 (0.9)	62 (2.3)	6 (0.2)
Type I		1247 (5.5)	16 (1.3)	16 (1.3)	54 (4.3)	1 (0.1)
Smoking history, n (%)						
Smoker		4816 (21.3)	19 (0.4)	29 (0.6)	106 (2.2)	9 (0.2)
Nonsmoker		17813 (78.7)	81 (0.5)	193 (1.1)	384 (2.2)	25 (0.1)
Steroid use, n (%)						
Steroid use		837 (3.7)	3 (0.4)	13 (1.6)	40 (4.8)	1 (0.1)
No steroid use		21792 (96.3)	97 (0.4)	209 (1.0)	450 (2.1)	33 (0.2)
History of bleeding disorder, n (%)		346 (1.5)	7 (2.0)	6 (1.7)	17 (4.9)	3 (0.9)
None		22283 (98.5)	93 (0.4)	216 (1.0)	473 (2.1)	31 (0.1)
Functional status, n (%)						
Dependent		489 (2.2)	1 (0.2)	11 (2.2)	20 (4.1)	3 (0.6)
Independent		22140 (97.8)	99 (0.4)	211 (1.0)	470 (2.1)	31 (0.1)
ASA score, n (%)						
3		10862 (48.0)	67 (0.6)	116 (1.1)	288 (2.7)	23 (0.2)
BMI, n (%)	30.6					

Feature	Average	Total	Cardiac Complication	VTE Complication	Wound Complication	Mortality
Comorbidities, n (%)						
Pulmonary		2253 (10.0)	12 (0.5)	21 (0.9)	61 (2.7)	4 (0.2)
Cardiac		12952 (57.2)	79 (0.6)	131 (1.0)	317 (2.4)	26 (0.2)
Complications, n (%)						
Mortality		34 (0.2)	18 (52.9)	3 (8.8)	0 (0.0)	34 (100.0)
Wound complications		490 (2.2)	2 (0.4)	14 (2.9)	490 (100.0)	0 (0.0)
VTE		222 (1.0)	9 (4.1)	222 (100.0)	14 (6.3)	3 (1.4)
Cardiac complications		100 (0.4)	100 (100.0)	9 (9.0)	2 (2.0)	18 (18.0)

ASA indicates American Society for Anesthesiology; BMI, body mass index; VTE, venous thromboembolism.

TABLE 2.

Area Under the Curve Values for LR, ANN, and ASA as Predictors of Cardiac Complications, VTE, Wound Complications, and Mortality

	LR (95% CI)		ANN (95% CI)		ASA (95% CI)	
Cardiac	0.657	(0.655–0.658)	0.710	(0.707–0.712)	0.468	(0.457–0.480)
VTE	0.588	(0.567–0.609)	0.567	(0.561–0.573)	0.435	(0.416–0.453)
Wound	0.613	(0.591–0.634)	0.606	(0.600–0.612)	0.491	(0.472–0.509)
Mortality	0.703	(0.682–0.724)	0.680	(0.673–0.686)	0.369	(0.352–0.386)

ANN indicates artificial neural networks; ASA, American Society for Anesthesiology; CI, confidence interval; VTE, venous thromboembolism.