A computational model to predict successful liberation from mechanical ventilation in ICU patients

Alexandra Szewc[1,3], Han Kim MSE[1,2,3], Robert D Stevens MD[2,3]

[1] Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA, [2] Anesthesiology and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, [3] Laboratory of Computational Intensive Care Medicine, Johns Hopkins University, Baltimore, MD, USA

Introduction

5-10% of planned extubations fail for all ICU patients.[1] Failed extubation is associated with significantly increased mortality and other adverse outcomes.[1] Mechanical ventilation (MV) is associated with many complications including pneumonia, lung injury, delirium, decreased physical activity, laryngeal damage and diaphragm dysfunction.[2] There is an unmet need for greater predictive accuracy to optimize decision-making and maximize successful extubation outcomes. Here, we explore machine learning models as a novel approach to predict outcomes after extubation in the ICU[3,4]

Objectives/Aims

A machine learning model that accurately predicts extubation success could serve as a decision support tool to identify the ideal duration of mechanical ventilation for each patient and to decrease the risk of complications. The aims of this study are to predict the likelihood of successful extubation and describe features that are associated with extubation outcome. We hypothesized that machine learning models could be trained to accurately and efficiently predict extubation outcomes.

Methods

The data used in this study was from the Philips eICU clinical research database which contains >200,000 ICU admissions from 208 institutions across the United States.[5] All adult patients who were extubated were included in the analysis. Patients were grouped into two classes: reintubated (n=7999) within 72h, and non-reintubated (n=30659). Features were considered from the entirety of a patient's first first time on MV with the task of learning physiology to predict whether the patient extubated would require reintubation. Exposure variables of interest included age, gender, laboratory results, MV duration, severity of respiratory failure, history of congestive heart failure, neurologic state, motor scores, sepsis/septic shock, fluid balance, medications, standard pulmonary variables, procedures, prior diagnoses, and physiology-derived data. The observation window included data from the 6 hours preceding extubation from each patient's first MV occurrence. The prediction window was 72 hours following extubation. Three different machine learning (ML) algorithms, generalized linear model (GLM), random forest (RF), and gradient boosting (XGboost), were evaluated.

Results

Model	XGBoost	Random Forest	GLM
Accuracy	0.71 (0.72, 0.70)	0.71 (0.74, 0.68)	0.60 (0.61, 0.60)
AUROC	0.81 (0.82, 0.80)	0.80 (0.81, 0.79)	0.70 (0.71, 0.69)
Sensitivity	0.78 (0.80, 0.76)	0.76 (0.79, 0.72)	0.72 (0.74, 0.70)
Specificity	0.70 (0.71, 0.68)	0.70 (0.74, 0.65)	0.57 (0.58, 0.56)
Precision Recall AUROC	0.52 (0.54, 0.51)	0.50 (0.53, 0.48)	0.35 (0.36, 0.34)
Precision	0.40 (0.41, 0.39)	0.40 (0.42, 0.37)	0.31 (0.31, 0.30)
Recall	0.78 (0.80, 0.76)	0.76 (0.79, 0.72)	0.72 (0.74, 0.70)

Table 1. Performance metric summary of the eICU-CRD developed model for the clinical endpoints of final extubation and extubation failure followed by reintubation. Results show a worse performance under the Generalized Linear Model, which is likely due to extraneous features that produce noise.

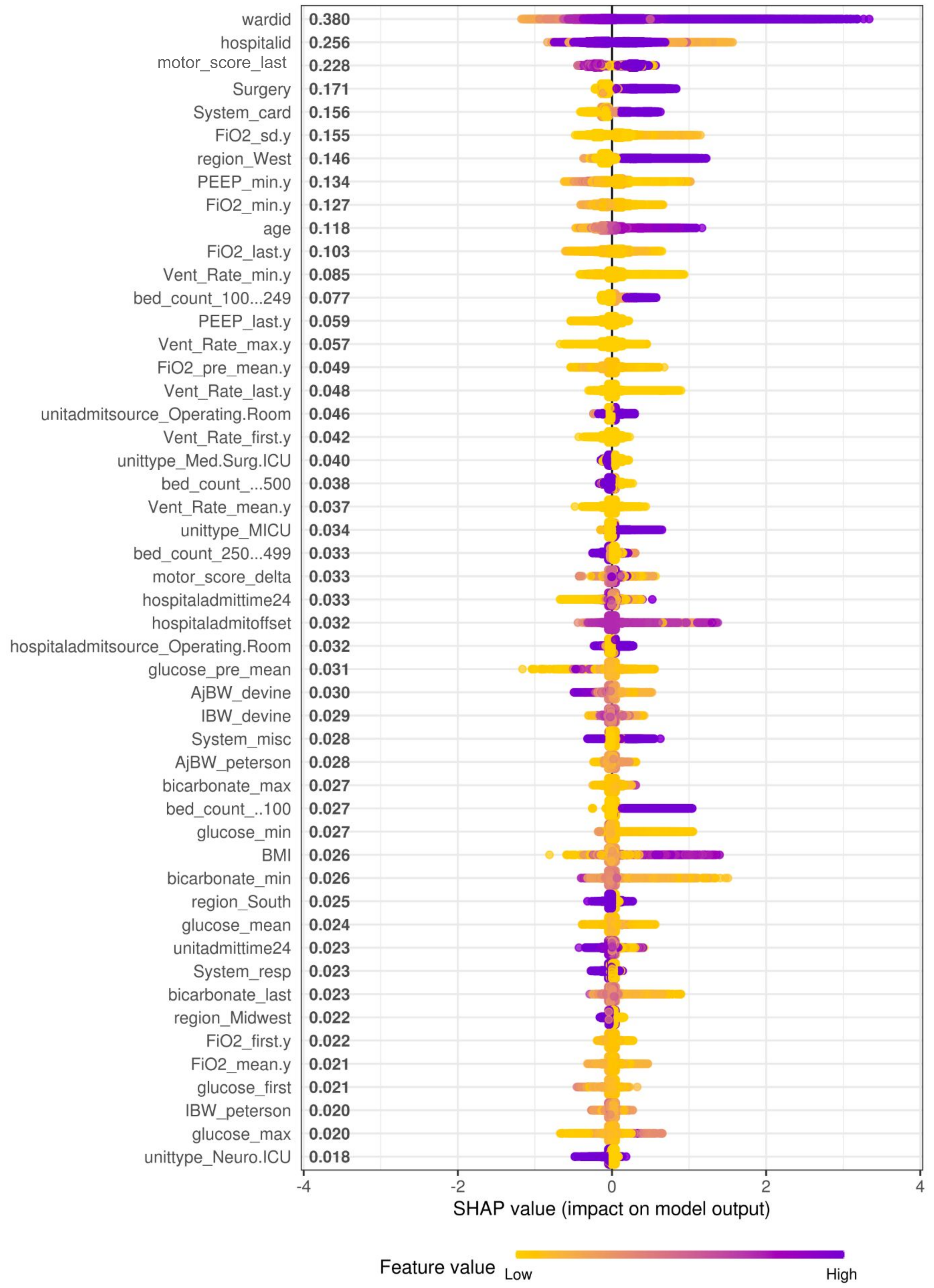

Figure 1 (top). XGBoost SHapley Additive exPlanations (SHAP) summary of the top 50 features over 5 outer fold iterations. Most important features are listed at the top. The color gradient represents the recorded value of each feature observation. A positive SHAP value indicates a higher probability of extubation failure.

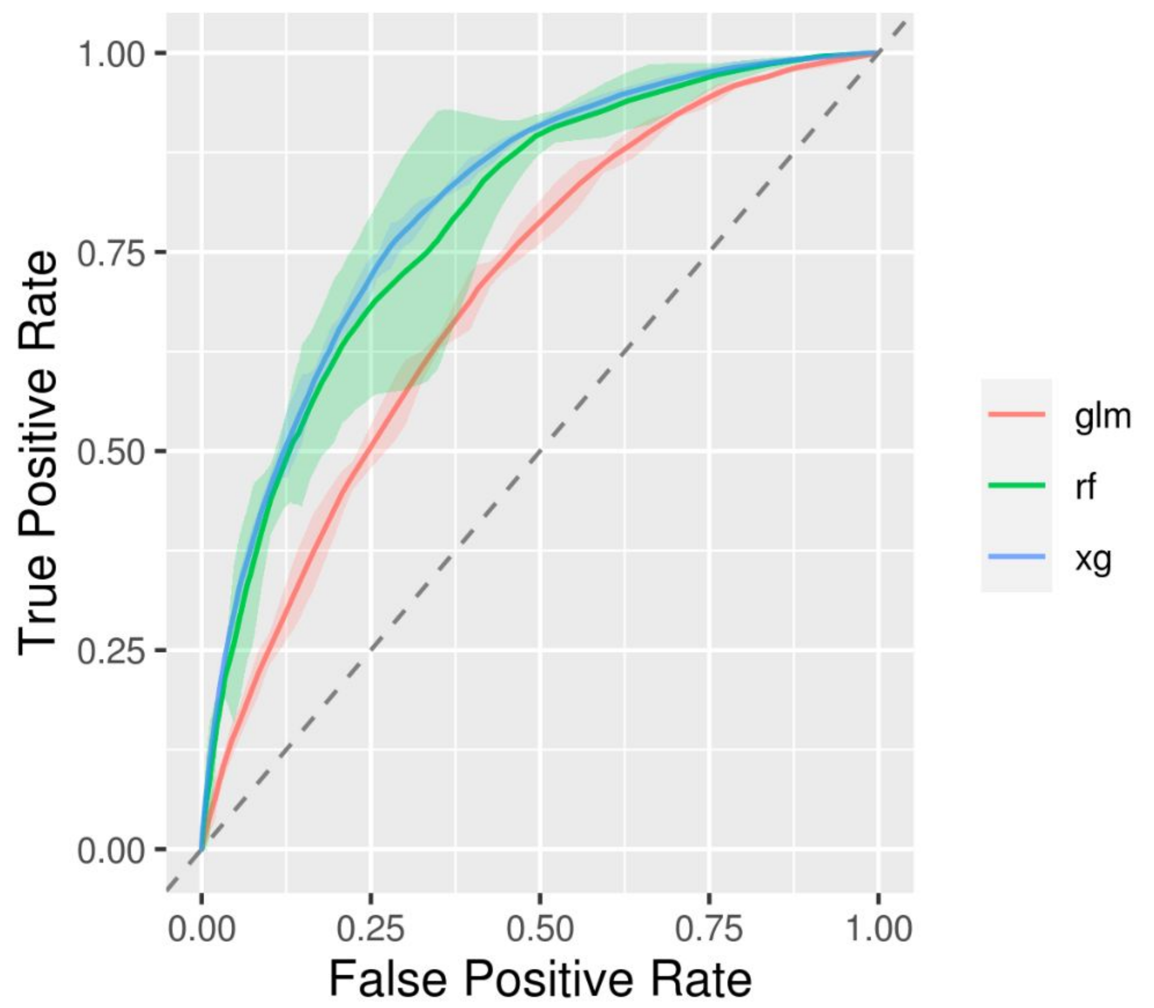

Figure 2 (left). Receiver operating characteristic curve and 95% confidence intervals of the generalized linear model (GLM), random forest (RF), and XGboost (XG) models. GLM models can be seen to struggle slightly due to increased dimensionality of the feature space. XGboost and Random forest models perform similarly with an average AUROC of 0.81 and 0.80 respectively.

Results (cont.)

Out of all the unique mechanically ventilated patients in the eICU database (n=38,769), the outcome class of non-reintubation (n=30,659) represented the majority of patients where the reintubation class (n=7,999) represented the remainder, excluding those who were reintubated following a time interval that exceeds 72 hours. XGBoost was the best performing model with AUROC: 0.81 ±0.01, sensitivity: 0.78±0.02, and specificity: 0.70±0.02.

Each model utilized a total of 114 features, which were selected from a collection of 1310 respiratory, demographic, clinical, and physiologic features. Of those features ranked within the top 50 as shown in Figure 2, the majority among the top 25 are MV or respiratory variables the eICU database. The most important features apart from these respiratory values include the last recorded motor score, surgical history, and age.

Conclusions and Continued Work

The findings indicate that there is significant potential for clinical data-driven machine learning approaches to serve as clinical decision support tools to aid in better preventing reintubation following mechanical ventilation in the intensive care unit. Our work demonstrates that computational models trained with ICU patient data recorded in the last 6 hours of mechanical ventilation can successfully predict extubation failure necessitating reintubation within 72 hours of extubation.

An important consideration for future improvements to these models will be the addition of external validation on the Medical Information Mart for Intensive Care (MIMIC) III database.[6] This will allow future results to be more generalizable and potentially applicable to real-world, clinical settings.

References

1. Thille AW, Richard JC, Brochard L. The decision to extubate in the intensive care unit. Am J Respir Crit Care Med. 2013;187(12):1294-1302. doi:10.1164/rccm.201208-1523CI
2. Goligher EC, Ferguson ND, Brochard LJ. Clinical challenges in mechanical ventilation. Lancet. 2016;387(10030):1856-1866. doi:10.1016/S0140-6736(16)30176-3
3. Meyer A, Zverinski D, Pfahringer B, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018;6(12):905-914. doi:10.1016/S2213-2600(18)30300-X
4. Tomašev N, Glorot X, Rae JW, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature. 2019;572(7767):116-119. doi:10.1038/s41586-019-1390-1
5. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data 5, 180178 (2018).
6. Johnson, A.E., et al. MIMIC-III, a freely accessible critical care database. Sci Data 3, 160035 (2016).