Novel Machine Learning Alert Model to Predict Cardiothoracic Intensive Care Unit Readmission or Mortality After Cardiothoracic Surgery

George A. Cortina, PhD [1,2], Shujin Zhong, BS[1], Marshall Nichols, MS[1], Michael Gao, MS[1], Will Ratliff, MBA[1], William Knechtle, MBA[1], Suresh Balu, MBA[1], Kelly Kester, MSN[3], Mary Lindsay, MSN[3], Jill Engel, MSN[3], Ashok Bhatta, MS[4], Jacob Schroder, MD[5], Ricardo Henao, PhD[6], Mark Sendak, MD[1], Mihai Podgoreanu, MD[4]

[1]Duke Institute for Health Innovation; [2]University of Virginia School of Medicine; [3]Duke University Health System/Heart Operations; Departments of [4]Anesthesiology, [5]Surgery, and [6]Electrical and Computer Engineering, Duke University.

Background.

Patients readmitted to the intensive care unit (ICU) after transfer to a lower acuity care unit demonstrate increased mortality, length of stay, and healthcare expenditures compared to those never readmitted. In high acuity specialized units, such as the cardiothoracic ICU (CTICU), identifying patients at highest risk of readmission and subsequent complications is challenging. Traditional rule-based risk tools overpredict complications, illustrating the need for dynamic means of prediction. Machine learning models have demonstrated improved success at identifying patients at risk for ICU readmission. However, these models use datasets from a variety of different ICU types (e.g. surgical vs. medical), which limits specificity. We sought to build a machine learning model that would run hourly in these intermediate care units to predict patient decompensation up to 48 hours before the event. We hypothesized that machine learning-based models can accurately monitor patients in intermediate care units to predict patient decompensation.

Methods.

A dataset was curated to include patient encounters between October 2015 and October 2018 that underwent cardiothoracic surgery then CTICU care and then transfer to an intermediate care unit. Input features included both patient summary data and hourly data. Patient summary data included past hospitalizations, comorbidities, and summary statistics of lab values from admission up to entry into the intermediate care unit. Hourly data included laboratory values along with provider actions such as orders and procedures while in the intermediate care unit assigned to each hour up until transfer from the intermediate unit or death. Hourly data included additional engineered features through cumulative summary statistics such as cumulative maximum, minimum, standard deviation, and mean values for numeric values. The primary outcome was an hour-by-hour based prediction of whether transfer back to the CTICU or death would occur in the next 48 hours. We evaluated the performance of several models including l1-penalized logistic regression and gradient-boosted decision trees. All models were evaluated against a held-out set consisting of 15% of the total data and hyperparameter settings were determined via a 3-fold cross-validation scheme. Final model performance was compared using both the area under the receiver operator characteristic (AUROC) and the average precision (AP). Hours within 48 hours of a patient's decompensation were considered positives while negative hours were ones outside of this window.

Results.

The identified cohort contained 5543 patients. Of these patients, 13 died and 549 had care escalated to a more intensive unit. All encounters comprised 732,879 total patient hours in the intermediate care unit. Of these hours, 16,267 or 2.22% were within 48 hours of an event. In total, there were 864 input features with 469 of these inputs being patient summary data and 395 hourly data features. Of the total 395 hourly inputs, 103 features were primary inputs with 202 engineered features. Of the 469 patient summary inputs, 260 inputs encoded past comorbidities, 175 inputs were engineered from 38 different lab values, and 34 were demographic features. Preliminary performance results indicate a total dataset AUROC/AP .70/.08 for l1-penalized logistic regression on the 15% test set. For gradient-boosted trees, performance results were .73/.11. Traditional score-based risk tools NEWS and MEWS scored .58/.03 and .57/.03 (AUROC/AP) respectively.

Conclusion.

We are developing a novel machine learning model designed to continually monitor patients for risk of decompensation in post-CTICU units. While preliminary, model performance suggests that we can discriminate between high-acuity patients that will decompensate and those that will not better than traditional score metrics. Training on hourly datasets allows for the creation of an alert model that ideally can run in a post-CTICU hospital unit on real time data to identify patients at risk of decompensation. Next steps will focus on improving model performance and then subsequent implementation. Accurate prediction of decompensation can be used to inform monitoring of patients and allow for action to avoid death or an emergent event.