Development of Machine Learning Model to Predict Risk of Inpatient Deterioration

Stephanie Skove, BS[1], Harvey Shi, BS[1], Ziyuan Shen, MS[2], Michael Gao, MS[1], Mengxuan Cui, MS[3], Marshall Nichols, MS[1], Suresh Balu, MS, MBA[1], Armando Bedoya, MMCI, MD[4], Dustin Tart, BSN[7], Benjamin A Goldstein, PhD[6], William Ratliff, MBA[1], Mark Sendak, MD, MPP[1], Cara O'Brien, MD[5]

[1]Duke Institute of Health Innovation, [2]Department of Electrical and Computer Engineering, Duke University, [3]Department of Biomedical Engineering, Duke University, [4]Division of Pulmonary, Allergy, and Critical Care Medicine, Duke University, [5]Department of Medicine, Duke University, [6]Department of Biostatistics & Bioinformatics, Duke University, Duke University Health System

Background.
The current practice for identifying deteriorating patients is reactive, rather than proactive. Patients often show early signs of deterioration hours before a rapid response team (RRT) or code blue is activated, and delays in care can have a detrimental impact on clinical outcomes. While it is vital that deterioration is identified early to prevent adverse outcomes, our ability to model and predict patient deterioration remains limited. Duke University's Department of Medicine and the Duke Institute for Health Innovation formed a transdisciplinary team to develop a machine learning model to accurately predict a patient's risk of deterioration.

Methods.
In this retrospective cohort study, data was curated from 171,120 adult patient encounters (age ≥ 18 at hospital admission) from three hospitals of a large academic health system from October 2015 to August 2018. We used a lightGBM (lgbm) model to obtain an hourly risk score to predict inpatient deterioration defined as an unanticipated ICU transfer or inpatient mortality within the next 24 hours. The training cohort was formed using a subset of all encounters representing 62,146 hospitalizations. 569 model features were built using 83 EHR data elements, including 18 vitals, 39 labs, 6 orders, and 17 medication administrations. In addition, for each analyte and vital sign measurement, a 24hour aggregate mean, minimum, and maximum were calculated along with the calculated change from prior measurement. The data was updated every hour and designed to estimate the probability of deterioration within the next 24hours. The model was tested on a separate cohort of 62,602 inpatient encounters from January 2019 to December 2019. Model performance was evaluated using the area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve (AUPRC) and compared to the National Early Warning Score (NEWS) model.

Results.
Of the 171,120 encounters 4.0% had the primary outcome (4,775 encounters with an unanticipated ICU transfer; 2,151 with an in-hospital mortality). The average age of patients experiencing a deterioration event was higher than those who did not deteriorate (64.7 vs 56.7). In addition, they spent an average of four extra days in the hospital (8.5 vs 4.7). On the 2019 test set, the lgbm model performed best with an hourly AUROC of 0.86 compared to the NEWS score with an AUROC of 0.78. The hourly AUPRC was 0.096 for the lgbm and 0.072 for NEWS.

Conclusions.
Initial development of a machine learning model accurately predicts the composite outcome of unplanned ICU transfer or inpatient mortality. Further research is needed to improve the sensitivity and precision of the model. Implementation of this model may improve care for patients at high risk of deterioration. Implementation of this model may improve care for patients at high risk of deterioration. The

interdisciplinary team aims to pilot this model in summer 2020 to support proactive rounding by rapid response teams.

.