TITLE

Development and Validation of Machine Learning Models to Predict Admission from the Emergency Department to Inpatient and Intensive Care Units

Alexander Fenn[1], Connor Davis[2], Neel Kapadia[3], Daniel Buckland[3], Marshall Nichols[2], Michael Gao[2], William Knechtle[2], Suresh Balu[2], Mark Sendak[2], B. Jason Theiling[3]

[1]Duke University School of Medicine, [2]Duke Institute of Health Innovation, [3]Division of Emergency Medicine, Department of Surgery, Duke University School of Medicine

Background.
This study aimed to develop and validate two machine learning models that utilize both historic and current-visit patient data from the EHR to predict the probability of patient admission to either an inpatient unit or ICU at each hour (up to 24 hours) of an ED encounter. Another aim was to provide a framework for clinical implementation of these machine learning models.

Methods.
Data was curated from 468,167 adult patient encounters to the 3 EDs of a large academic health system from August 1st, 2015 to October 31st, 2018. Of these encounters, 84.9% of patients were discharged, 13.5% were admitted to an inpatient ward, and 1.7% were admitted to an ICU. The following features were built: age, chief complaint, mode of arrival, and Emergency Severity Index (ESI) from ED intake; vitals, orders, medication administrations, and analyte results during the ED encounter; and comorbidities and admissions and hospital utilization within the year preceding the ED encounter. A gradient-boosted tree model (Light GBM) was developed and evaluated at each hour (up to 24 hours) of the patient's ED encounter, beginning with preadmission (using historical patient data, chief complaint, and ESI score), and then incorporating additional data from initial triage presentation to data collected throughout the patient's encounter. The most recent 10% of data was withheld for a test set. From the remaining 90% of data, 10% of encounters were pulled for training and 10% for validation. Both AUROC and average precision score were used to evaluate model performance. The model was then validated using a cohort with identical inclusion criteria as the initial cohort, using data from January 1st, 2019 to December 31st, 2019.

Results.
The AUROC for the intermediate admission model was 0.873, and the AUPRC was 0.636. For the ICU admission model, the AUROC was 0.951, and the average precision score was 0.461. The five most important features for the intermediate model were: Patient age, WBC count, hematocrit level, cumulative length of stay of visits to the ED resulting in discharge in the past 365 days, and platelet count. For the ICU admission model, the five most important features were: Patient age, hematocrit level, WBC count, cumulative length of stay of visits to the ED resulting in discharge in the past 365 days, and ESI score.

For the 2019 validation data, the AUROC for the intermediate admission model was 0.854, and the AUPRC was 0.692. For the ICU admission model, the AUROC was 0.834, and the AUPRC was 0.224. When model performance was measured by hour, for the intermediate admission model, the AUROC was highest between 0 and 2 hours, and the average precision score increased throughout the encounter time, reaching its max value at 23 hours. For the ICU admission model, the AUROC and average precision score were highest between 0 to 2 hours, and then decreased after that time period.

Conclusion.
Machine learning models can be developed to accurately make predictions regarding the probability of inpatient or ICU admission throughout the entire duration of a patient's encounter in the ED, not just at

the time of triage. These models remain accurate on a patient cohort beyond the time period of the initial training data and can be integrated to run on live EHR data with similar performance.