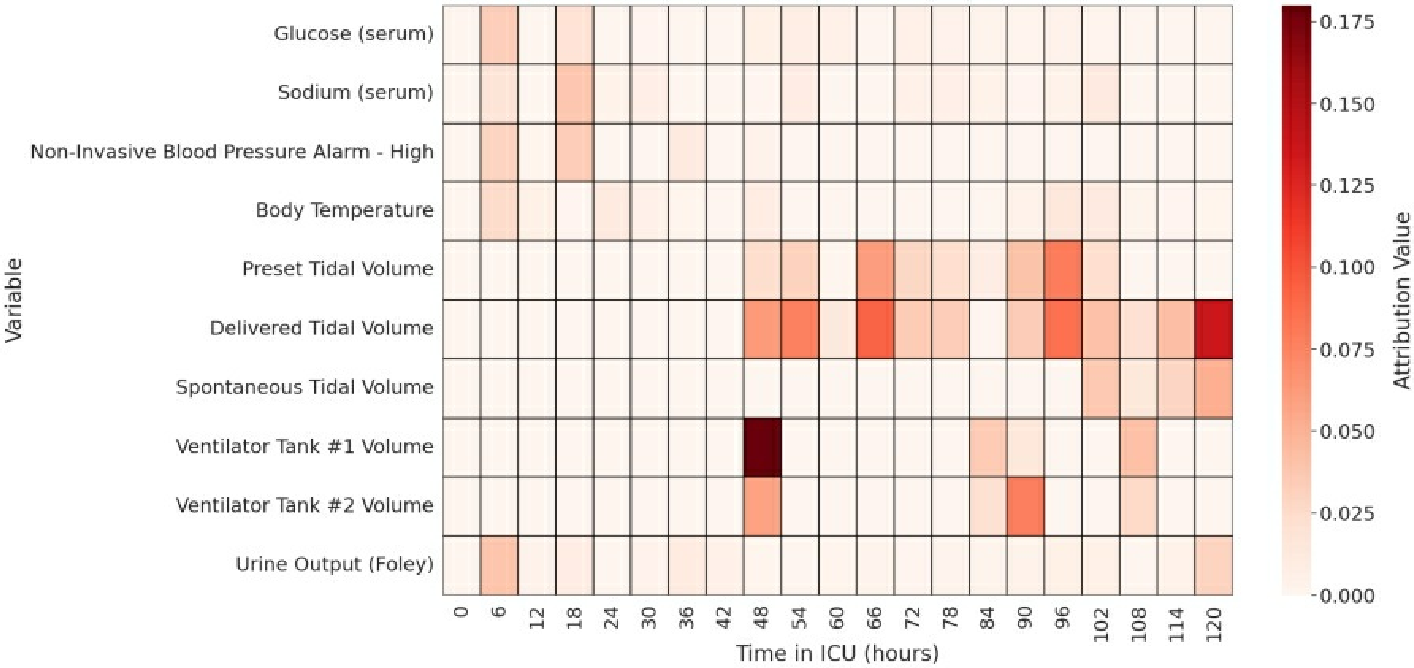
Time-series Machine Learning Approach to Sepsis Prediction in the Intensive Care Unit

I. Sears[1], M. M. Levy[2], C. E. Ventetuolo[3], C. Eickhoff[4], A. Abbasi[3]; [1]Warren Alpert Medical School, Brown University, Providence, RI, United States, [2]Brown University/Rhode Island Hosp, Providence, RI, United States, [3]Medicine, Brown University, Providence, RI, United States, [4]Center for Biomedical Informatics, Brown University, Providence, RI, United States.

Corresponding author's email: isaac_sears@brown.edu

Rationale: Early detection of sepsis has proven an attractive target for applications of machine learning. We combined a novel time-series machine learning model with an interpretability approach to develop a model to predict sepsis onset in the intensive care unit (ICU), identify variables important to model performance, and determine if variable importance changes over time. Methods: We applied a Time-Series Transformer (TST) to electronic health record data within the Medical Information Mart for Intensive Care database version four (MIMIC-IV) to predict sepsis (defined by Sepsis-3 criteria) six hours before its onset. ICU stays that met Sepsis-3 criteria within the first 24 hours, representing subjects likely admitted with a diagnosis of sepsis, were excluded. We then applied Captum, a deep learning interpretability tool, to generate “attributions” to explain the TST’s predictions. For each data point, the magnitude of the corresponding attribution indicated how important that data point was to the final prediction. Attributions corresponding to individual stays were used to generate heatmaps to depict the change in variable importance for sepsis prediction over the course of the stay (Figure). Attributions were aggregated across all stays and time points to determine the variables most important for predicting sepsis onset. Polynomial regression models were fit to the attributions aggregated over ICU stays, but not time, to determine whether there was statistically significant change in attribution over time. Results: Our study cohort included 31,733 ICU stays. Subjects had a mean age of 64.1 years ± 17.0. Approximately half of the subjects were men (55.0%), a majority were white (67.8%), and 2,504 (7.9%) subjects developed sepsis. Mean length of stay was 76.8 hours ± 79.2 and the mean time to sepsis onset was 81.6 hours ± 21.6. The mortality rate was 7.2%. The variables most important for predicting sepsis included urine output, ventilator settings, vital signs, and blood glucose levels. The importance of these variables changed over the course of the stay, with significant fluctuations in attribution over time. Conclusion: The TST applied to unprocessed electronic health record data performed well at predicting sepsis onset in the ICU. Captum can be used to identify variables important for the TST’s performance and the change in variable importance over time. Clinical predictive models that do not take time course into account may misrepresent the variables important to sepsis prediction.
Figure: Attribution heatmap for a single ICU stay. In this example, ventilator settings were most important for predicting sepsis.



This abstract is funded by: National Institutes of Health grant U54 GM115677 (I. S), R01 HL141268 (CEV), P20 GM 103652 and Parker B. Francis Foundation (A.A.).