

## Predicting ICU length of stay with a competing risk deep learning survival model framework

Xinggong Liu, MD, PhD<sup>1</sup> and Omar Badawi, PharmD, MPH<sup>1,2</sup>

<sup>1</sup> Patient Care Analytics, Philips Healthcare, Baltimore, Maryland; <sup>2</sup> Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts

**Background.** Intensive care unit (ICU) length of stay (LOS) is often used as a quality outcome to measure ICU efficiency, with the target being lower than expected LOS. However, it has been well established that this approach can penalize ICUs with a high quality of care as very high-risk patients may die very quickly in their ICU course. However, with better ICU care they may survive but often with a prolonged LOS. Therefore, ICUs which increase the survival of high-risk patients will have longer than expected LOS performance. Unfortunately, traditional ICU prediction models do not account for this potential bias. We hypothesize that ICU LOS as a quality measure can be improved by accounting for the competing risks of being discharged alive or deceased.

We implemented the DeepHit (Lee, C., Zame, W.R., Yoon, J., & Van der Schaar, M (2018)) approach, a deep learning survival model with competing risks to predict ICU LOS accounting for the two competing events of being discharged alive or deceased. In addition to the utility of predicting both the time to death and time to discharge alive simultaneously, we also hypothesized this neural network model would gain performance over traditional, regression-based models by not making any assumptions about the relationship between risk factors and survival time distribution.

**Methods.** ICU stays with an ICU LOS greater than 24 hours were extracted from the Philips eICU Research Institute (eRI) database between 2012 and 2016. The sample was randomly split into a development and validation set at an 8:1 ratio. Because the ICU mortality rate was relatively rare (4.6%), surviving patients were under-sampled in the development set to force a ratio of 1:2 for the deceased and surviving patients. The DeepHit neural network of a shared sub-network and two cause-specific subnetworks was trained to learn the joint distribution of the two competing events (deceased and survived), with the goal of minimizing the log-likelihood of the combined joint distribution of the first occurring event and the sum of the cause-specific ranking loss function. Model features included the following available at 24 hours after admission: baseline patient characteristics, including basic demographics, vital signs, laboratory measurement, and hourly ICU mortality risk (a separate generalized additive model trained in the development set with ICU mortality as the outcome). The model hyper-parameters were selected based at three evaluation times of 16, 36 and 75 hours of follow-up (based on the actual distribution of LOS after 24 hours of ICU admission: lower, middle and upper quartile). Model performance was evaluated using a time-dependent concordance index at each of the three evaluation times. When comparing model performance using APACHE IVa as a reference, the median predicted time to first occurring event was used as the LOS prediction and compared with APACHE IVa ICU LOS prediction in the Root Mean Square Error (RMSE).

**Results.** The final sample included 38,478 patients stays for development and 32,341 stays for validation. The DeepHit model that produced the highest performance metric consisted of a four-layer network with one fully connected layers for the shared sub-network and four fully connected layers for the two events sub-network. The C-indexes at 16, 36 and 75 hours of index time were (0.845, 0.778), (0.813, 0.765) and (0.787, 0.724) for time to death and time to discharge alive, respectively. When compared with APACHE IVa ICU LOS prediction, which only produces a single LOS prediction without mortality status, the new model's LOS prediction (time to first occurring event) exhibited higher accuracy through a lower RMSE (87.3 vs 89.1) compared to APACHE.

**Conclusion.** We have demonstrated that a DeepHit neural network model that learns the joint survival distribution of two competing events would offer the added utility of providing ICU LOS prediction conditioned on mortality status. This approach will be useful for quality benchmarking to mitigate the penalty assigned to ICUs with many high-risk survivors.