Computational Subphenotype Discovery and Validation of ICU Stratum Traumatic Brain Injury Patients

Han Kim MSE[1,2], Robert D Stevens MD[1,2]

[1] Department of Anesthesiology and Critical Care Medicine, [2] Laboratory of Computational Intensive Care Medicine
Johns Hopkins University School of Medicine, Baltimore, MD, USA

Introduction

Traumatic Brain Injury (TBI) is a significant health hazard worldwide that not only displays high incidences of mortality (approximately 20% of TBI cases lead to death) but also ultimately results in life-long cognitive deficits and motor dysfunctions in many patients.[1] Despite the progress in clinical advancements to detect the severity of TBI through CT/MRI imaging and notable identification biochemical derangements such as perturbations of homeostasis, increased free radical generation, inflammation apoptosis, and diffuse axonal injury, to date, there has yet to be any promising clinical trials to further advance TBI treatment.[5] The leading consensus is that there is significant pathophysiological heterogeneity within the TBI patients and that each phenotype of TBI exhibits varying responses to treatment.[6] Therefore, heterogeneity within TBI populations is recognized as a major barrier in efforts to find effective treatments and improve outcomes.

Objectives/Aims

The overarching premise for this work is that **existing paradigms do not capture the complexity of TBI** which encompasses a broad array of clinical and biological features. **We hypothesize that combinations of features extracted from clinical electronic health records (EHR) and from physiological time series (PTS) monitoring data can be segregated using unsupervised machine learning, enabling discovery of latent, data-driven subphenotypes that have distinct likelihoods of clinical outcomes.** The results of this study have potential to significantly enhance the ability to differentiate TBI patients based on quantifiable pathophysiological information, leading to better treatment selection and increased efficacy of developed treatments catering to a specific phenotype.

Methods

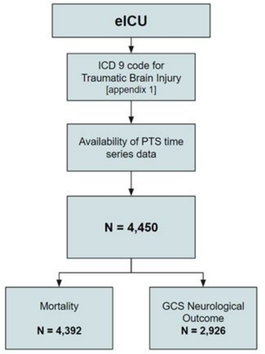


Figure 1. Study inclusion and exclusion criteria identifying the 4,450 eICU TBI patients, 346 MIMIC III TBI patients were identified following the same criteria. Based on the availability of clinical endpoints, the final modeled sample size differed.

	eICU (n = 4,450)	MIMIC (n = 346)
Age (sd)	61.57 (21.9)	61.70 (21.77)
Gender = male (%)	2,706 (60.8)	207 (59.8)
Mild TBI (%)	2,340 (52.6)	158 (45.7)
Moderate TBI (%)	1,081 (24.3)	71 (20.5)
Severe TBI (%)	1,029 (23.1)	117 (33.8)
Neurological outcome (mGCS) = unfavorable outcome (%)	657 (22.5)	55 (16.9)
In hospital mortality = expired (%)	546 (12.4)	40 (11.6)

Table 1. Demographic Summary of the TBI population in the eICU-CRD and MIMIC-III database.

Adult TBI patients (N=4,450) were identified in a multi-center ICU database (eICU) and clinical, laboratory and PTS data were extracted. Statistical PTS features were derived from heart rate, SaO2, blood pressure, and respiratory rate high frequency bedside signals. Unsupervised clustering algorithms were applied accounting for mixed data types on a statistically pruned 147 derived variables. The discovered clusters were then characterized according to outcomes at discharge, and differences in physiology, then externally validated on TBI patients in the independent MIMIC III dataset. A multiclass classification model was trained using the identified clusters as endpoints/labels. This allowed us to classify the MIMIC III TBI patients (346) into one of the eICU identified clusters.

Results

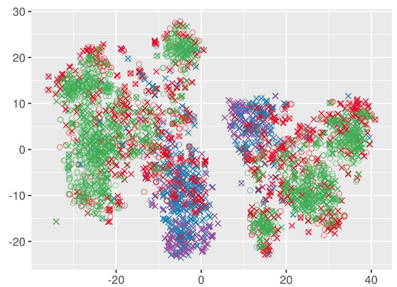


Figure 2. t-SNE dimensionality reduce plot showing t-SNE dimension 1 vs dimension 2 overlaid with colors corresponding to clusters and "x" and "o" for clinical outcome. Similar to a PCA representation of our 147 features to represent our TBI cohort, the t-SNE plot is noisy but able to present some level of visible clusters identified via unsupervised learning method.

We identified four TBI clusters (a, b, c, d) each with a distinct outcome probability distribution, and each associated with a unique, clinically relevant pattern of PTS and laboratory features. Subphenotype (a) captures TBI patients whose physiological features are associated with the highest likelihood of survival and favorable neurological outcome, while subphenotype (d) captured patients whose physiological features are associated with the highest risk of death and unfavorable neurological outcome, while subphenotypes (b) and (c) had intermediate outcome probabilities. Both the physiologic and outcome differences between clusters were reproduced in the MIMIC III cohort when eICU clusters were assigned using a multi-class classification. The mortality and neurological outcome proportions per subphenotype for the eICU cohort and assigned MIMIC III cohort can be seen in Figure 3.

Figure 4 is a heatmap visualizing the differences between subphenotypes. The physiologic differences per patient between clusters show distinct data-driven physiologic signatures unique to each cluster. To provide supporting evidence as to how these unique signatures transferred to clusters assigned to MIMIC III, Figure 5 shows the feature comparison between subphenotypes (a) and (d) compared between clusters identified in eICU and assigned in MIMIC III.

Figure 4. Heatmap of the normalized top 40 discriminative feature values for all four subphenotypes (a, b, c, and d) the eICU TBI cohort. The difference between the four subphenotypes can be visually seen in the heatmap with each subphenotype characterized by a range of physiologic value that defines as we see in figure 3 a specific illness severity and probability of clinical outcome.

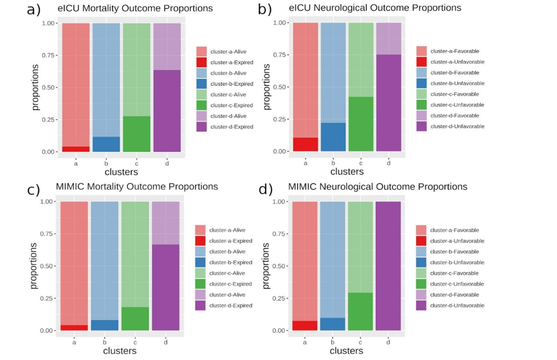
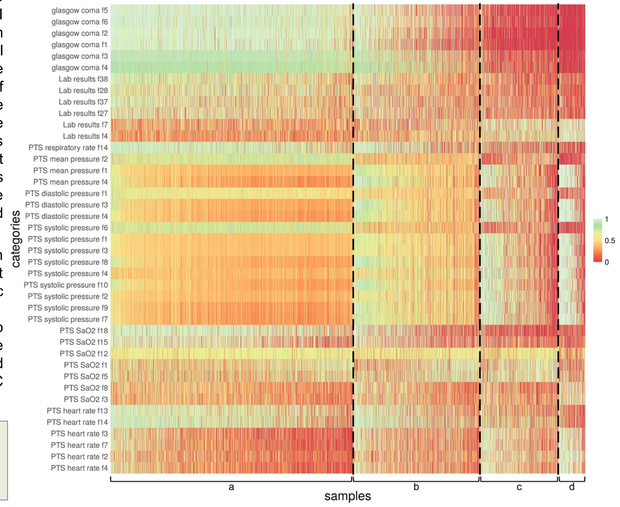


Figure 3. Mortality and Neurological Outcome Proportions corresponding to subphenotypes discovered in eICU (a, b) and externally validated in MIMIC III (c, d). MIMIC III cohorts classified using an eICU trained multiclass model show that assigning clusters based on EHR and PTS derived features can successfully categorize TBI patients based on eICU identified outcome proportions and illness severity.



Results (cont.)

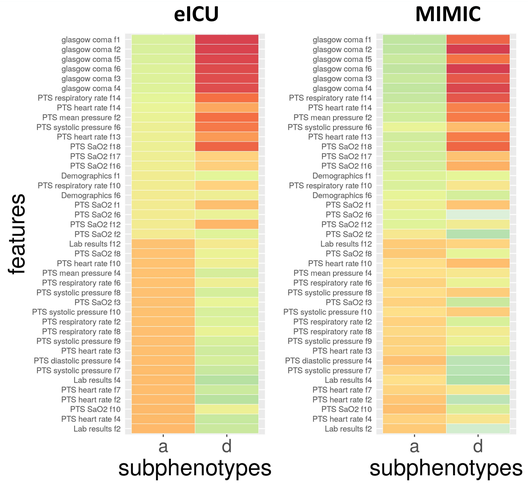


Figure 5. Standardized mean differences (SMD) of the top 40 discriminative features represented as heatmaps comparing subphenotype a and d as they correspond in Figure 1. (left) eICU SMD showing differences in individual variables and the standardized degree to which subphenotypes a and d differ. (right) MIMIC III SMD between subphenotypes a and d resulting from assigning MIMIC III TBI cohort to eICU identified subphenotypes. We can clearly observe similar magnitude of standardized mean differences between clusters a and d for eICU identified and MIMIC III assigned TBI cohorts.

Conclusions & Future Work

Using unsupervised machine learning applied to EHR and PTS derived features of TBI patients admitted to the ICU, we identified four distinct and clinically meaningful TBI clusters. Patients assigned to specific clusters had distinct outcome probabilities and unique data-driven physiologic signatures which suggest that they are plausible candidate subphenotypes. Results indicate a novel approach to categorizing ICU stratum TBI patients based on objective, numerical patient physiological and metabolic data. Moreover, the same four eICU TBI subphenotypes were successfully validated in the MIMIC III TBI cohort. Ongoing research will explore other characteristics of these TBI subphenotypes and in particular their differential response to specific treatments and interventions as well as extend the analyses to MIMIC IV, Amsterdam University Medical Center (AUMC), and Johns Hopkins TBI cohorts.

References

1. Stocchetti, N., et al., Traumatic brain injury in an aging population. J Neurotrauma, 2012. 29(6): p. 1119-25.
2. Dutton, R.P. and M. McCunn, Traumatic brain injury. Curr Opin Crit Care, 2003. 9(6): p. 503-9.
3. Perel, P., et al., Systematic review of prognostic models in traumatic brain injury. BMC Med Inform Decis Mak, 2006. 6: p. 38.
4. Maas, A.I.R., et al., Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. Lancet Neurol, 2017. 16(12): p. 987-1048.
5. Dijkland, S.A., et al., Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies. J Neurotrauma, 2019.
6. Volovici, V., et al., Intensive care admission criteria for traumatic brain injury patients across Europe. J Crit Care, 2019. 49: p. 158-161.