8

Limited value of Dynamic Physiological data in Ordinal Outcome Prediction in moderate to severe traumatic brain injury

David Bark[1], Ludvig Hult[2], Anders Hånell[1], Teodor Svedung-Wettervik[1], Anders Lewén[1], Per Enblad[1], Thomas Schön[2], Elham Rostami[1,3]
[1] Division of Neurosurgery, Department of Medical Sciences, Uppsala University, Uppsala, Sweden , [2] Department of Information Technology, Uppsala University, Uppsala, Sweden, [3] Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden

Background: Refined outcome prediction of the full Glasgow outcome scale extended (GOSE) in TBI patients could improve risk stratification compared to binary outcome. However, predicting the exact GOSE category is challenging due to the complex pathophysiology of TBI. Dynamic variables from the intensive care unit (ICU) may provide critical information regarding secondary insults to the brain and have the potential to improve outcome predictions beyond static admission variables.

Methods: In a single center cohort from the years 2007-2024 (n=516) we developed XGBoost and ordinal logistic regression models comparing temporal (by year of trauma) and stratified (by GOSE outcome distribution) train-test split when predicting 7-grade GOSE (1-2 combined) at 24h and 48h post-injury. We have conducted extensive comparisons of models using only baseline clinical data versus models augmented with a comprehensive set of features derived from high-frequency ICU data.

Results: Across both 24h and 48h times, models incorporating dynamic ICU features performed equally or worse than models using only baseline variables. For example, the Quadratic Weighted Kappa (QWK) for the 24-hour Ordinal Logistic Regression model was 0.48 (95% CI: 0.28 - 0.64) with baseline features and 0.48 (95% CI: 0.29 - 0.63) after adding ICU features. Critically, model performance consistently degraded on a temporal test split compared to the more commonly used stratified split.

Conclusion: Contrary to expectations, extensive feature engineering of physiological data did not improve ordinal outcome prediction. This suggests the prognostic signal is either already captured by robust baseline variables or is confounded by unmeasured treatment intensity. Furthermore, model performance dropped significantly on a temporal test split, highlighting challenges for clinical implementation of these types of models during changes in patient populations and clinical practices over time.