

# Early Prediction of Length of Stay in Hospitalized Patients with Stroke and Traumatic Brain Injury

Heramb Gupta<sup>1</sup>; Darsh Patel<sup>1</sup>; Shreya Hari<sup>1</sup>; Steven Solar<sup>1</sup>; Nihao Sun<sup>1</sup>; Runtian Tang<sup>1</sup>; J. L. Greenstein<sup>1</sup>; C. O. Taylor<sup>1</sup>; R. D. Stevens<sup>2</sup>

<sup>1</sup>Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

<sup>2</sup>Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

## Background

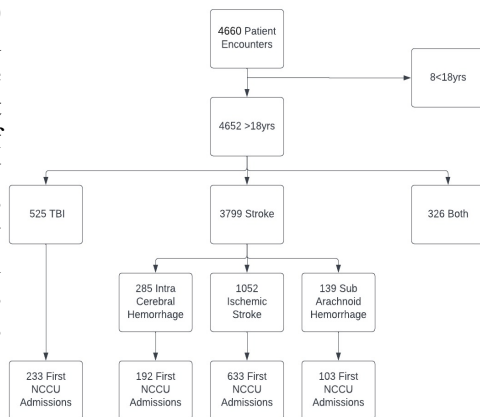
**Stroke** is one of the leading causes of morbidity and mortality worldwide, and **traumatic brain injury (TBI)** is one of the major causes of disability in children and young adults. **ICU length of stay (LoS)** is considered a **primary driver of inpatient costs**. The prediction of length of stay in the early phase of hospitalization can inform resource allocation and improve clinical decision-making to ultimately reduce medical spending

## Objectives

- To use patient **data available in the first 24 hours of stay** to predict length of stay (LOS) for patients with traumatic brain injury and stroke in the Neuro Critical Care Unit (NCCU).
- To identify and rank predictive features driving the LOS.

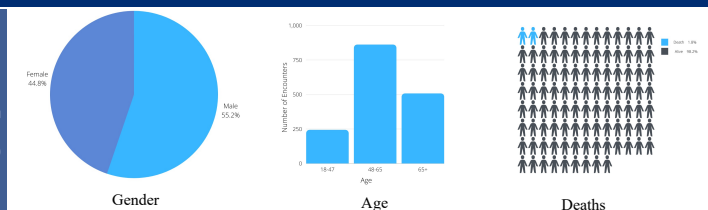
## Methods

The **JHU Precision Medicine Analytics Platform (PMAP)** database was searched for adults admitted to the Neuro Critical Care Unit for management of stroke or TBI. Study subject selection process is shown in the flow diagram. Model performance was evaluated with measures of discrimination, precision, and calibration



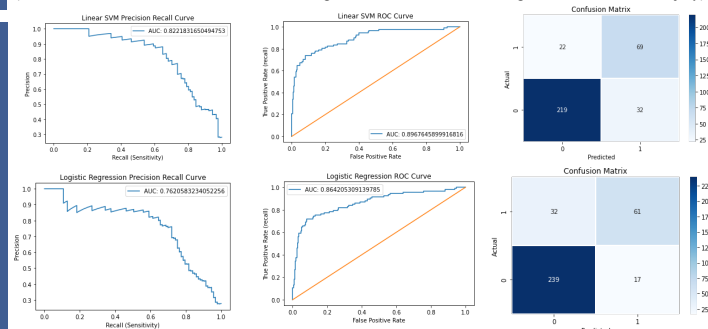
## Results

**Patient Demographics**



**21 Patients (1.8%) died in their first 24hours in the NCCU** (9-Ischemic Stroke, 2-Sub-Arachnoid Hemorrhage, 1-Intra Cerebral Hemorrhage, 9-Traumatic Brain Injury)

**LOS Model for Entire Patient Cohort**



Parameters	SVM	LR
AU –ROC	0.897	0.864
AU – PR	0.822	0.762
F1 Score	0.842	0.860

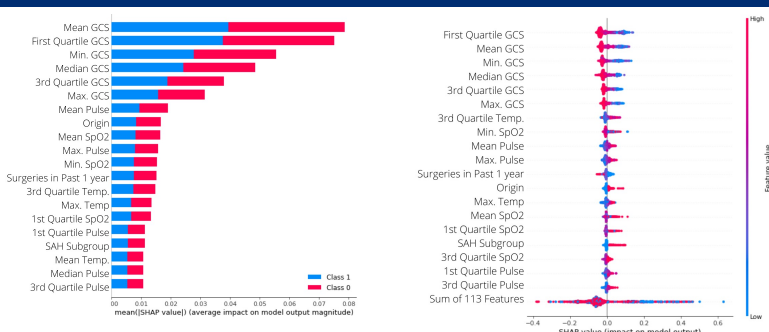
Support Vector Machine outperforms Logistic Regression

## Input Features and Output Labels

- There were 132 input features consisting of Clinical and physiologic data available in the first 24h after admission which were used to train different ML classifiers to predict LOS (output variable)
- Patients were classified as a 1 if their LOS was in the highest quartile (or death) and 0 otherwise.

## Results

**Feature Importance**



The **Glasgow Coma Scale (GCS)** is the most important feature in predicting NCCU Length of Stay.

**Different LOS Models**

SVM Performance for different patient groups

	Entire Cohort	TBI	Stroke	Ischemic Stroke	Intra Cerebral Hemorrhage	Sub Arachnoid Hemorrhage
No. of Features	132	115	139	136	124	121
Cutoff (in days)	4.60	3.06	5.03	3.90	6.01	14.78
AUC ROC	0.897	0.852	0.830	0.715	0.900	0.500
AUC PR	0.822	0.780	0.756	0.629	0.733	0.607
F1 Score	0.842	0.895	0.795	0.769	0.864	0.786

The **best performing model –Support Vector Machine** is that for the **entire patient cohort** which has information about TBI and stroke types as input features.

## Conclusions

**Accurate predictions about length of stay for NCCU** can be made using **patient data available in the first 24 hours**. Considering an average cost of 4000\$ per day in the NCCU, cost of stay can be estimated

## Future Scope

The models will be validated on external available ICU datasets. Quantification of length of stay may correlate with NCCU cost, and the complex relationship between the two may be quantified further.