



OPEN

## Rapid prediction of secondary neurologic decline after traumatic brain injury: a data analytic approach

Jamie Podell<sup>1,2</sup>, Shiming Yang<sup>1,3,4</sup>, Serenity Miller<sup>1</sup>, Ryan Felix<sup>1</sup>, Hemantkumar Tripathi<sup>1</sup>, Gunjan Parikh<sup>1,2</sup>, Catriona Miller<sup>1</sup>, Hegang Chen<sup>1,4</sup>, Yi-Mei Kuo<sup>1</sup>, Chien Yu Lin<sup>1</sup>, Peter Hu<sup>1,3,4</sup> & Neeraj Badjatia<sup>1,2</sup>✉

Secondary neurologic decline (ND) after traumatic brain injury (TBI) is independently associated with outcome, but robust predictors of ND are lacking. In this retrospective analysis of consecutive isolated TBI admissions to the R. Adams Cowley Shock Trauma Center between November 2015 and June 2018, we aimed to develop a triage decision support tool to quantify risk for early ND. Three machine learning models based on clinical, physiologic, or combined characteristics from the first hour of hospital resuscitation were created. Among 905 TBI cases, 165 (18%) experienced one or more ND events (130 clinical, 51 neurosurgical, and 54 radiographic) within 48 h of presentation. In the prediction of ND, the clinical plus physiologic data model performed similarly to the physiologic only model, with concordance indices of 0.85 (0.824–0.877) and 0.84 (0.812–0.868), respectively. Both outperformed the clinical only model, which had a concordance index of 0.72 (0.688–0.759). This preliminary work suggests that a data-driven approach utilizing physiologic and basic clinical data from the first hour of resuscitation after TBI has the potential to serve as a decision support tool for clinicians seeking to identify patients at high or low risk for ND.

Traumatic brain injury (TBI) accounts for over two and a half million emergency department visits in the United States annually<sup>1</sup> and is a leading cause for evacuation from austere military and civilian environments<sup>2,3</sup>. While many individuals recover fully, TBI results in enormous costs to society as the second leading cause of disability in the United States<sup>1</sup>. Acute trauma care has long stressed the importance of rapid assessment and skilled treatment during the “golden hour” of resuscitation, when interventions are most likely to prevent long-term morbidity and mortality<sup>4</sup>. Advances in point-of-care tools have aided in the rapidity by which systemic trauma patients can be prioritized, as with the Extended Focused Assessment with Sonography in Trauma (E-FAST) exam<sup>5</sup>; however, without analogous brain injury risk stratification tools, early triage and management of TBI patients remains challenging, especially in resource-limited austere environments<sup>3</sup>.

The primary triage concern, regardless of initial severity, is determining the likelihood of secondary neurological decline (ND), defined by clinical or radiographic worsening or need for neurosurgical intervention. An estimated 20% of moderate to severe<sup>6</sup> and 5–10% of mild<sup>7</sup> TBI patients experience ND, which is independently associated with worse outcome<sup>8,9</sup>. Appropriately timed medical and surgical interventions may prevent or mitigate secondary brain injury associated with ND, whereas prolonged monitoring and routine follow-up imaging may consume unnecessary resources in low-risk patients. Therefore, identifying individuals at highest and lowest risk of ND is a primary goal in initial TBI triage and resuscitation.

However, during the first hour of TBI resuscitation, risk stratification is based off of rudimentary data including the neurologic exam, static vital signs (VS), and clinical judgment<sup>9</sup>. While none have translated to standard clinical practice, there is growing interest in point-of-care technologies including ultrasound, pupillometry, and blood biomarkers for early TBI triage<sup>3,10,11</sup>. Tools utilizing dynamic VS for prediction of ND are particularly attractive given ubiquitous VS monitoring across settings. Further, non-specific early warning scores using static

<sup>1</sup>Program in Trauma, Shock Trauma Neurocritical Care, University of Maryland School of Medicine, 22 S. Greene Street, G7K19, Baltimore, MD 21201, USA. <sup>2</sup>Department of Neurology, University of Maryland School of Medicine, Baltimore, USA. <sup>3</sup>Department of Anesthesiology, University of Maryland School of Medicine, Baltimore, USA. <sup>4</sup>Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, USA. ✉email: nbadjatia@som.umaryland.edu

VS to predict clinical deterioration already have shown some utility in TBI patients<sup>12</sup>. More nuanced physiologic markers and trends—detectable by machine learning (ML) but not routine clinical observation—may better characterize and predict patient trajectories during the golden hour of TBI resuscitation by providing insight into autonomic nervous system dysfunction associated with impending ND. ML is emerging as a useful tool to predict deterioration and need for intervention in numerous clinical scenarios<sup>13,14</sup>, including trauma resuscitation<sup>15,16</sup>. While IMPACT and CRASH models can aid in TBI prognostication by using early admission data to predict mortality and 6 month outcome<sup>17,18</sup>, there remains a critical need to predict more proximate potentially actionable events that might contribute to longer term outcomes. This is the goal of our work aiming to predict early ND, to support early triage and clinical decision making.

We previously demonstrated that analysis of photoplethysmography (PPG) and electrocardiography (ECG) waveform and variability data during the first 15 min of resuscitation could better predict ND than clinical variables alone in a predominantly mild TBI cohort<sup>19</sup>. In the present study, we expand upon these findings in a larger TBI cohort with the goal of identifying physiological markers within the first hour of resuscitation, from data accessible to pre-clinical or austere environments, to aid in the clinical determination of ND risk over the subsequent 48 h.

## Materials and methods

**Patient selection, study design, and data acquisition.** This is a retrospective cohort study of consecutive TBI patients cared for in the Trauma Resuscitation Unit at the R Adams Cowley Shock Trauma Center at the University of Maryland Medical Center. Included patients met the following criteria: 1. Diagnosis of survivable TBI [Head Abbreviated Injury Score(AIS) 1–5]<sup>20</sup>; 2. Direct admission from scene of injury (while a reliable exact time from incident to admission was not available for these patients, the majority arrive within one hour of injury); 3. Age > 18 years old. 4. No major systemic injuries (Thoracic or Abdominal AIS  $\leq$  1); 5. No confounding active substance abuse (toxicology screen negative for cocaine and opiates). Patients were then excluded for the following: 1. Hospital length of stay less than 48 h; 2. Missing or insufficient physiologic data, defined as < 20 min of the first hour recorded. Baseline demographic and clinical data was obtained from the institutional trauma registry.

Continuous physiologic data was obtained from ECG, PPG, and arterial blood pressure (ABP) waveforms collected via BedMaster® (Excel Medical Electronics Inc., Jupiter, FL) VS collection system for 60 min beginning at the time of hospital arrival<sup>21</sup>. Trended VS such as heart rate (HR), peripheral capillary oxygen saturation (SpO<sub>2</sub>), and respiratory rate (RR) were collected every two seconds (0.5 Hz), and ECG, PPG, and ABP waveforms were collected at 240 Hz. All head computed tomography (CT) scans performed within the first 48 h were independently rated using Rotterdam and Marshall scoring systems<sup>22,23</sup> by study investigators (NB, GP, JP), who were blinded to demographic and outcome data.

This study was performed as part of the Real-Time Vital Sign Assessment to Predict Neurological Decline After Traumatic Brain Injury (RAPID-TBI) study, funded by the Department of Defense (FA8650-18-2-6H18). The study was approved by the institutional review board of the University of Maryland, Baltimore (HP-00060944), which waived the need for informed consent. All research was performed in accordance with the Declaration of Helsinki<sup>24</sup> and the Health Insurance Portability and Accountability Act<sup>25</sup>.

**Outcome measures.** The primary outcome was ND between one and 48 h of admission. As in previous work<sup>19</sup>, ND was a binary variable defined by the occurrence of one or more clinical or radiographic worsening events. Clinical ND was defined by a spontaneous decrease in Glasgow Coma Scale (GCS) by 2 or more points, loss of pupillary reactivity, development of pupillary asymmetry of  $\geq$  2 mm, treatment of raised intracranial pressure, or need for neurosurgical intervention. Hourly clinical exam findings were obtained from review of electronic medical record (EMR) (EPIC Systems Cooperation) nursing flowsheets. Each case of clinical ND was reviewed by study investigators to ensure clinical decline was not attributable to use of sedatives, analgesics, neuromuscular blockade, or dilated ophthalmologic exams. Clinical ND events were subclassified as neurosurgical ND when progressive cerebral edema resulted in a neurosurgical intervention, the timing of which was ascertained via EMR review of operative notes. These clinical ND events were adjudicated via consensus among reviewing investigators (NB, GP, JP).

Radiographic ND was defined by a  $\geq$  1 point worsening of the Rotterdam<sup>23</sup> severity score from initial to follow-up CT. While often insensitive to findings such as hemorrhage expansion, the Rotterdam is reliably associated with outcome after TBI<sup>23,26</sup>, and was therefore selected over other more granular radiographic features to ensure the clinical relevance of radiographic ND. As part of routine institutional practice all patients with moderate to severe TBI underwent follow up head CT imaging at 6 and 24 h after admission. Patients presenting with mild TBI (GCS 13–15) underwent CT imaging based on findings of initial scan and/or presenting mechanism of injury. Interrater reliability of Rotterdam scores assessed by three authors was calculated using data from 180 patients (approximately 20% of 905 total patients). The Cronbach alpha (95% confidence interval) was 0.966 (0.949, 0.979), and there was 100% interrater agreement regarding cases of radiographic ND (where Rotterdam worsened by  $\geq$  1 point).

If a patient experienced more than one ND event within the first 48 h, then time of ND was defined by the earliest event. Individuals who experienced ND within the first hour of admission were excluded from analysis given desire to use the first hour of data to forecast future ND. Additional secondary outcomes were collected via review of trauma registry data. These included hospital length of stay, mortality, and discharge destination.

Descriptive and univariate statistics were performed comparing presenting demographic, clinical, and injury characteristics and hospital outcomes across patients with and without ND. The Wilcoxon rank sum test was

used for ordinal variables and the Chi-squared test for categorical variables using a threshold of  $p < 0.05$  for statistical significance.

**Physiologic data processing.** High-fidelity PPG and ECG waveform and trended VS data collected during the first hour of admission was pre-processed by removal of extreme values deemed outside of normal physiologic ranges (HR > 200 bpm, SBP > 250 mmHg, DBP > 200 mmHg) and extreme outliers (based on a moving median with a window of 30 s) before extraction of features<sup>27</sup>. The ECG and PPG signals were smoothed with a robust discretized spline smooth filter to improve the signal-to-noise ratio<sup>28</sup>. These filtering steps may be automatically applied in a real time prospective manner. As previously described<sup>15</sup>, VS features including mean, standard deviation (SD), median, interquartile range (IQR), dose above and below thresholds, and first, second, and third quartiles of HR, systolic blood pressure (SBP), diastolic blood pressure (DBP), shock index (SI = HR/SBP), and SpO<sub>2</sub> were extracted from trended VS data. PPG and ECG-based heart rate variability (HRV) time and frequency domain features were calculated using standard definitions based on the Task Force of the European Society of Pacing and Electrophysiology<sup>29</sup> (Supplemental Table 1). Frequency domain features were calculated using three distinct popular methods, each with their own strengths and weaknesses<sup>30</sup>, including Welch's method of averaging periodograms from overlapping intervals<sup>31</sup>, analysis of least-squares based Lomb periodograms<sup>32</sup>, and parametric autoregressive modeling<sup>33</sup>. Non-linear dynamics HRV features included measures of entropy, Poincare plots, and fractal analyses<sup>34</sup>.

Our prior work suggested that clinical measurements, such as admission GCS, could improve prediction of ND<sup>19</sup>. However, adding these variables requires manual evaluation with clinical expertise. To build a fully automated physiologic model, we created GCS and Injury Severity Score (ISS)<sup>20</sup> estimation models that use only VS variables, derived from a set of non-overlapping patient data collected prior to this study<sup>35,36</sup>. These models were used to generate additional variable inputs for ND prediction models, namely, model outputs for the prediction of specific GCS and ISS scores<sup>36</sup>. The GCS estimation model generates the probability that a patient's GCS at admission is a particular value from 3 to 15. The ISS estimation model outputs the probability that a patient's ISS falls into one of five categories: 1 (1–4), 2 (5–8), 3 (9–15), 4 (16–24), and 5 (25–75). This is a form of transfer learning<sup>37</sup>, where lessons learned from large-scale physiologic data predictions are transferred as a substitution for the value being predicted into a new model.

**Prediction model development.** Three ND prediction models were created based on different kinds of predictor variables (clinical, physiologic, and combined models). The clinical model utilized presenting demographic information (age, sex, race/ethnicity), injury characteristics (type, mechanism), and clinical exam (arrival GCS as documented by TRU staff), all of which would have been known within the first hour of hospital presentation. The physiologic model utilized VS features from the first hour of admission, including the transfer learning variables described above. The combined model included both clinical and physiologic features from the first hour of admission.

To study the time-to-event data and handle model nonlinearity and missing values, the gradient boosted survival tree (GBS) model was used<sup>38</sup>. The GBT uses an ensemble of decision trees to sequentially learn the error left over by the previous tree. It has been demonstrated to outperform other machine learning methods in CPH analysis, such as random forest or support vector machine<sup>39</sup>.

To simplify model training and improve interpretation (e.g. variable importance), we used the Maximum Relevance Minimum Redundancy (MRMR) method to select features<sup>40</sup>. This method attempts to find a subset of features that maximize the association while minimizing the redundancy (measured by mutual information) between the features and the outcome. We selected 20 features that satisfied the MRMR.

During training, to prevent over-fitting, we tuned model hyperparameters including total number of trees, maximum tree depth, learning rate, and percentage of sampled variables. Through fivefold cross-validation, optimal hyperparameters were identified via random grid search. SHapley Additive exPlanations (SHAP) values were calculated and displayed graphically in order to describe each variable's contribution to ND prediction<sup>41</sup> and improve model interpretability<sup>39</sup>.

Model performance was described using time-dependent receiver operating characteristic (ROC) areas under the curve (AUC) and their 95% confidence intervals (CI). Overall (time-independent) performance for each model was described by its concordance index (C-index) and 95% CI<sup>42,43</sup>. Model development and statistical analyses were performed by Extreme Gradient Boosting (XGBoost), an efficient Python implementation of GBT (xgboost, v1.3.3)<sup>44</sup>. Missing waveform data was missing completely at random<sup>45</sup> and occurred in 20 patients (approximately 2%). It was handled by imputing with the most frequent values.

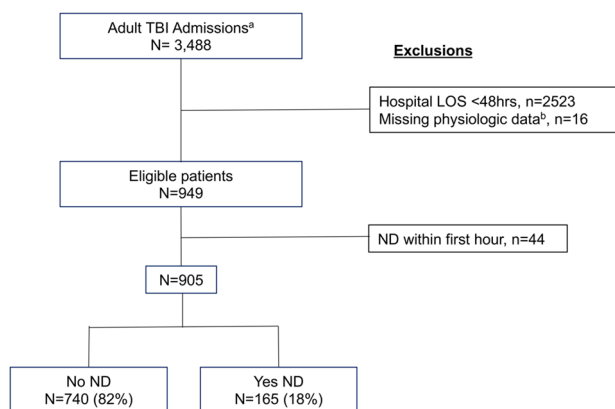
As an exploratory analysis, we assessed whether our ND prediction model could also predict in-hospital mortality. To this end, we used the ROC analysis to evaluate the association between ND model output and in-hospital mortality. This model's performance was described by the AUROC and its 95% CI.

## Results

**Baseline characteristics.** Admission characteristics of patients grouped by primary outcome are described in Table 1. Of 905 eligible TBI cases, 165 (18%) experienced ND of any kind within 48 h of admission (Fig. 1). The type and timing of ND events are displayed in Fig. 2; clinical ND occurred in 130 (79% with ND; 14% overall), radiographic ND in 54 (33% with ND; 6% overall), and neurosurgical ND in 51 patients (31% with ND; 6% overall). More than one ND type occurred in 57 patients (35% with ND; 6% overall), with all three types occurring in 13 (8% with ND; 1% overall). The median time to ND was 5 h (IQR 3–10). Several arrival VS demonstrated small but statistically significant associations with ND, including temperature and oxygen saturation.

	Neurologic decline		p-value
	Yes	No	
	N = 165 (18)	N = 740 (82)	
Age <sup>1</sup>	57 (38, 75)	54 (33, 71)	0.099
Sex <sup>2</sup>			0.242
Male	116 (70%)	485 (66%)	
Female	49 (30%)	255 (34%)	
Injury type <sup>2</sup>			< 0.001
Blunt	147 (89%)	707 (96%)	
Penetrating	18 (11%)	33 (4%)	
Arrival GCS <sup>1</sup>	13 (7, 15)	15 (13, 15)	< 0.001
Arrival vitals <sup>1</sup>			
Temp	36.6 (36.3, 36.8)	36.7 (36.5, 36.8)	< 0.001
SBP	150 (133, 177)	146 (130, 177)	0.599
DBP	89 (75, 100)	89 (77, 100)	0.783
HR	86 (73, 100)	89 (77, 102)	0.079
RR	22 (18, 27)	22 (18, 26)	0.377
O2Sat	99 (96, 100)	98 (95, 100)	<b>0.037</b>
Race <sup>2</sup>			0.462
White	97 (59%)	450 (61%)	
Black	48 (29%)	227 (31%)	
Asian	5 (3%)	10 (1%)	
Other/Unknown	15 (9%)	53 (7%)	

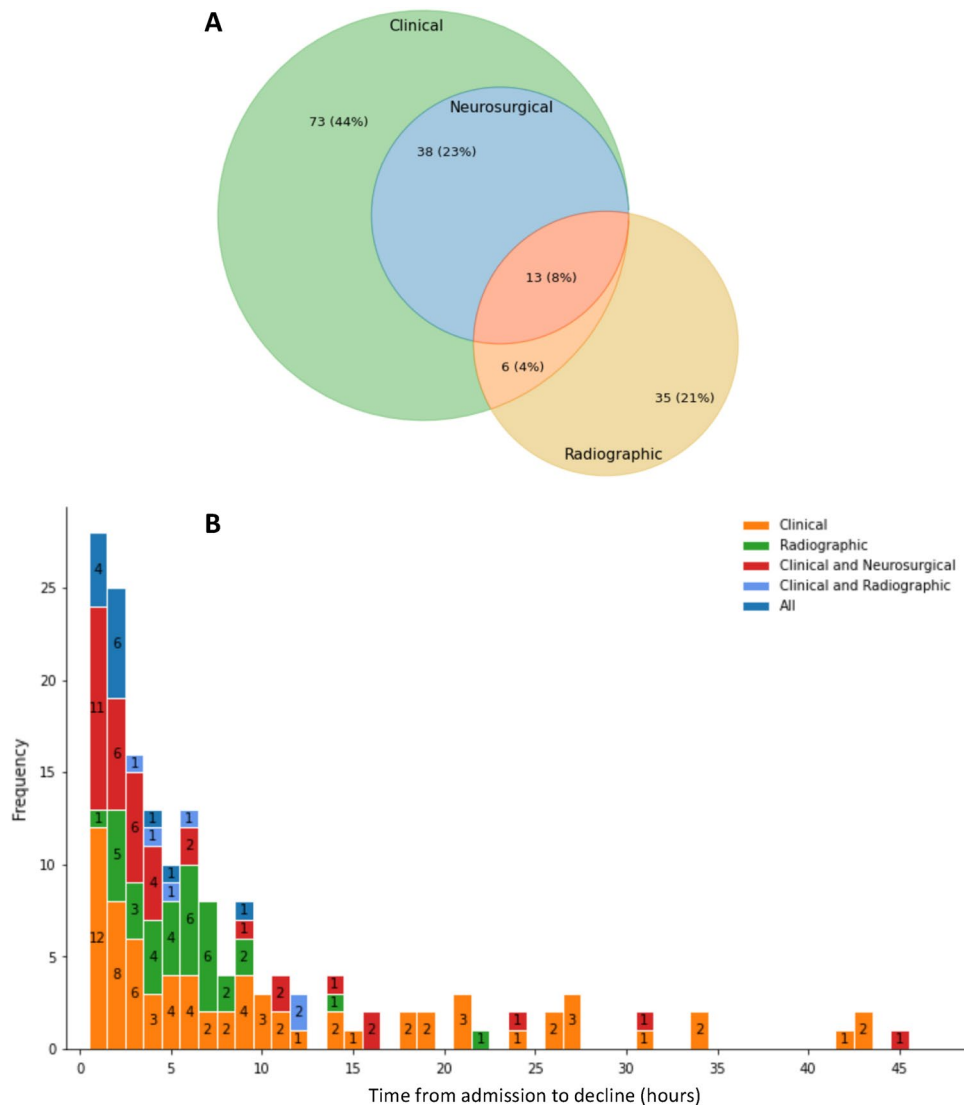
**Table 1.** Demographic, clinical, and injury characteristics. Statistical significance and p-values were assessed for ordinal variables using Wilcoxon rank sum tests<sup>1</sup> and for categorical variables using Chi-squared tests<sup>2</sup>. GCS = Glasgow coma scale; Temp = temperature (degrees Celcius); SBP = systolic blood pressure (mmHg); DBP = diastolic blood pressure (mmHg); HR = heart rate (beats per minute); RR = respiratory rate (breaths per minute); O2Sat = oxygen saturation (%). Significant values are in bold.



**Figure 1.** <sup>a</sup>All cases identified from trauma registry with Head AIS 1–5, Thoracic/Abdominal AIS ≤ 1, and toxicology screen negative for opiates and cocaine, from Nov 2015–Jun 2018. <sup>b</sup>Cases included if ≥ 30% of first hour continuous physiologic data was available for analysis. TBI = traumatic brain injury; LOS = hospital length of stay; ND = neurologic decline.

Clinical factors found to be associated with ND on univariate analysis included penetrating injury ( $P < 0.001$ ) and lower GCS ( $P < 0.001$ ).

Inpatient clinical characteristics and hospital discharge outcomes of patients grouped by primary outcome are described in Table 2. Patients who developed ND had a higher incidence of invasive ICP monitoring, higher in-hospital mortality, lower discharge rates to home, and longer hospital lengths of stay ( $p < 0.001$ ). Initial head CT findings are reported in Supplemental Table 2.



**Figure 2.** A. Ven diagram displaying the prevalence of ND events by subtype. B. Temporal distribution of ND, with y-axis representing the number of patients experiencing ND during each hour, as denoted by the x-axis, color-coded by ND subtype.

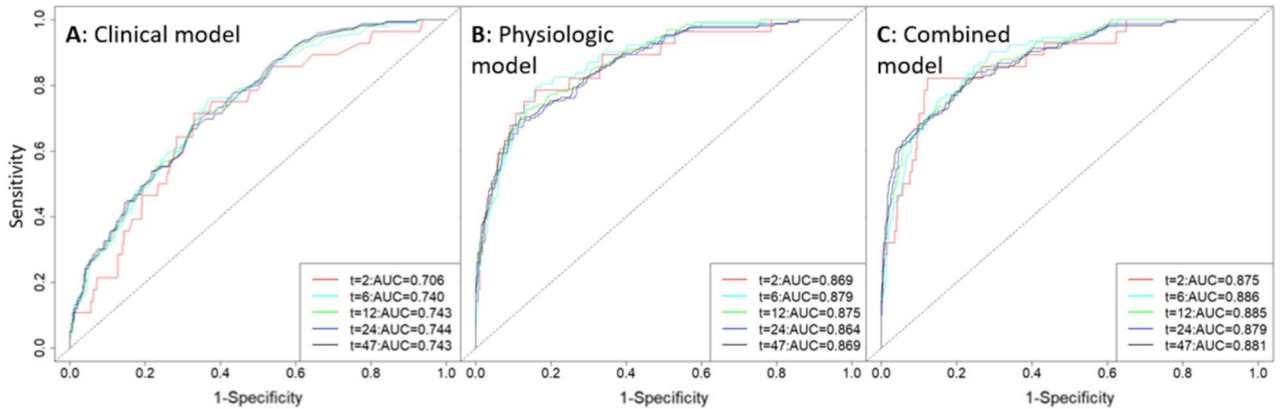
**Model performance and description.** CPH model performance for prediction of ND based on clinical only, physiologic only, and combined features is graphically displayed in Fig. 3. Time-dependent AUCs and 95% confidence intervals are displayed graphically in Fig. 4. The C-index of the physiological model was 0.84 (95%CI: 0.81–0.87), while the C-index of the clinical model was 0.72 (95%CI: 0.69–0.76). The combined model performed similarly to the physiological model, with C-index of 0.85 (95%CI: 0.82–0.88). Our supplemental mortality analysis demonstrated an AUROC of 0.79 (95%CI: 0.73–0.85) in the prediction of in-hospital mortality based on combined ND model output (Supplemental Fig. 1). Contributing feature SHAP values for models are displayed in Fig. 5. In this figure, features are ordered by importance along the y axis, with the most important contributors to the model at the top. Each colored dot represents a single patient's feature value, with red hue representing higher values and blue hue representing lower values, arranged along the x-axis, according to the log odds ratio that the variable contributed to the model prediction.

## Discussion

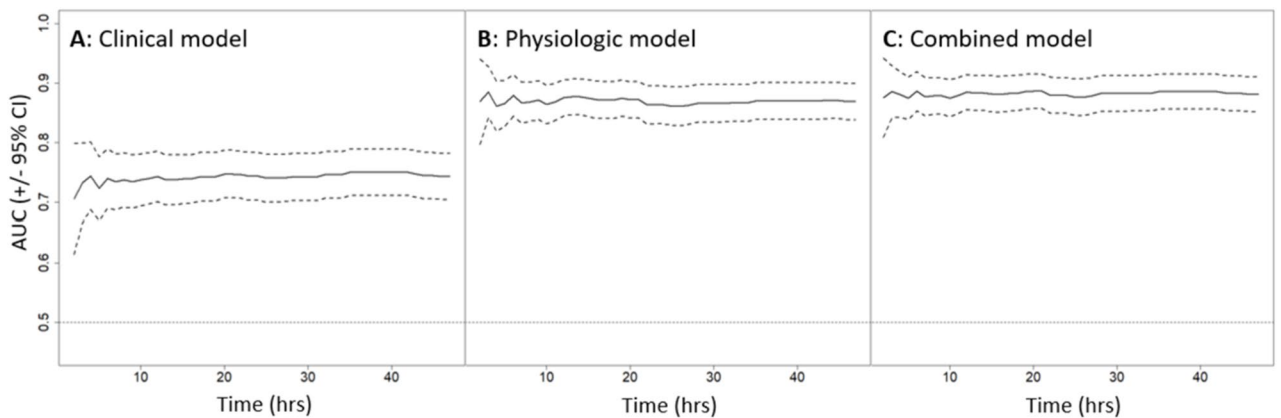
In this analysis of 905 TBI patients, we demonstrated that a ML algorithm based on data gathered from the first hour of hospital resuscitation was able to discern time to secondary decline within the first 48 h of admission. As expected, ND was a common clinically meaningful event associated with mortality and worse in-hospital outcomes. This analysis provides early proof of concept that utilizing continuous VS data to develop an automated early warning score for acute TBI patients undergoing inpatient observation may be feasible and has the potential to improve timely triage and resource utilization.

	Neurological decline (N = 905)		p-value
	Yes ND, N = 165 (18%)	No ND, N = 740 (82%)	
<b>Admission destination<sup>2</sup></b>			
ICU	67 (41%)	112 (15%)	<b>&lt; 0.001</b>
IMC	42 (25%)	271 (37%)	
OR	54 (33%)	253 (34%)	
Floor	0 (0%)	68 (9%)	
None/DC	2 (1%)	36 (5%)	
<b>ICP monitoring<sup>2</sup></b>	34 (21%)	43 (6%)	<b>&lt; 0.001</b>
<b>Survivor discharge disposition<sup>2</sup></b>			
Home	23 (18%)	286 (40%)	<b>&lt; 0.001</b>
Facility	107 (82%)	427 (60%)	
<b>Mortality<sup>2</sup></b>	35 (21%)	27 (4%)	<b>&lt; 0.001</b>
<b>Hospital LOS (days)<sup>1</sup></b>	9.1 (5.5, 17.0)	5.2 (3.3, 9.0)	<b>&lt; 0.001</b>

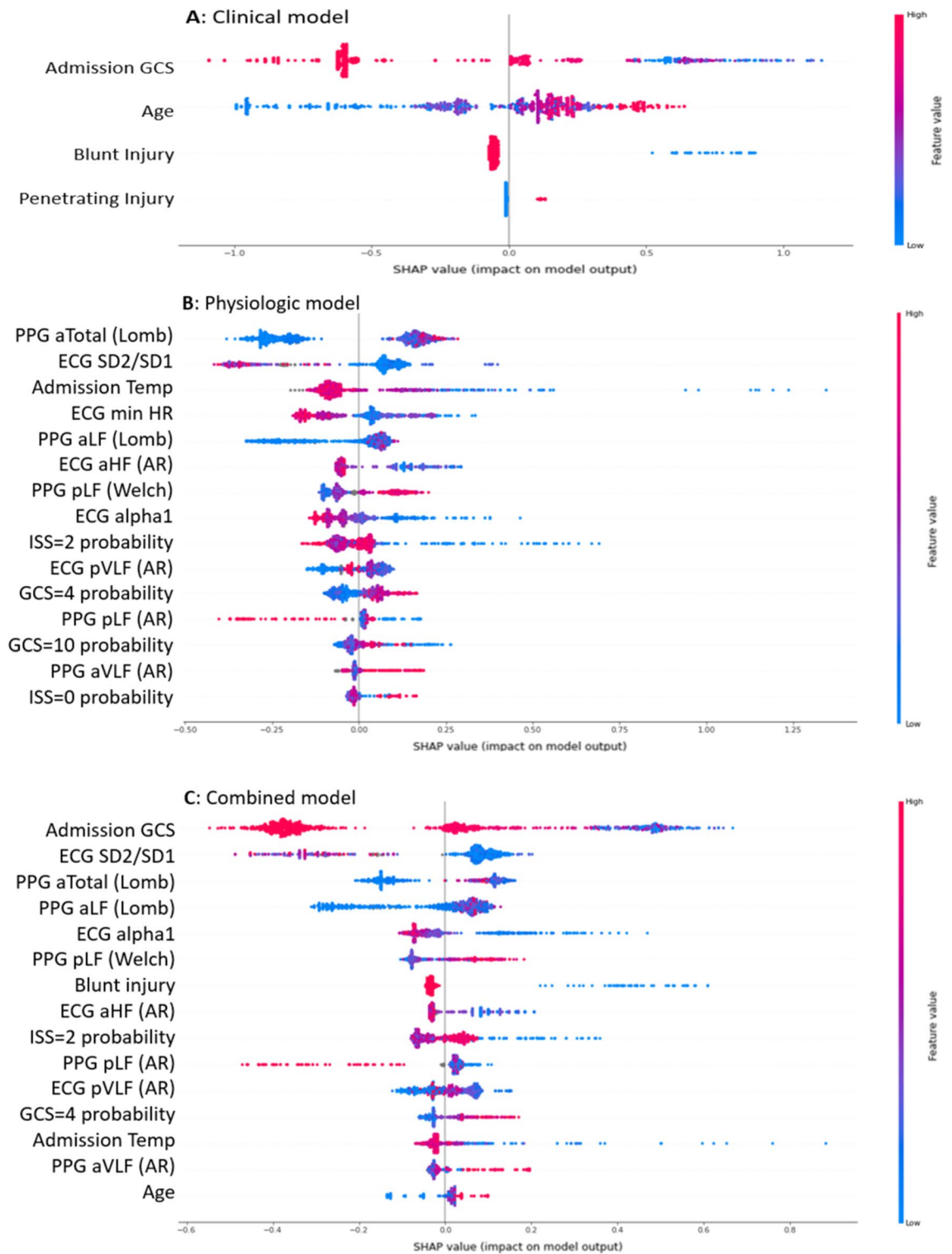
**Table 2.** Clinical characteristics and outcomes of hospitalization. Statistical significance and p-values were assessed for ordinal variables using Wilcoxon rank sum tests<sup>1</sup> and for categorical variables using Chi-squared tests<sup>2</sup>. ICP = intracranial pressure; LOS = length of stay; ICU = intensive care unit; IMC = intermediate care unit; OR = operating room; DC = discharged. Significant values are in bold.



**Figure 3.** Neurologic decline (ND) prediction model performance based on clinical (A), physiologic (B), and combined (C) predictor variables. Individual receiver operating characteristic (ROC) curves demonstrate model performance for predicting ND at specific times (in hours) from presentation, denoted by line color.



**Figure 4.** Time-dependent areas under the curve (AUC) of receiver operating characteristic (ROC) analysis (solid lines) displayed with 95% confidence intervals (CI, dotted lines) for ND prediction models based on clinical (A), physiologic (B), and combined (C) predictor variables. The time-independent concordance index (C-Index) for each model was 0.72 (95%CI: 0.69–0.76), 0.84 (95%CI: 0.81–0.87), and 0.85 (95%CI: 0.82–0.88), respectively.



**Figure 5.** For the clinical (A), physiologic (B), and combined (C) neurologic decline (ND) prediction models, each contributing feature is displayed on the y axis. The x axis shows each feature’s Shapley Additive Explanations (SHAP) values. A larger SHAP value denotes a higher log odds ratio that a variable’s value added to the prediction. Values are represented in color ranging from red to blue (high to low). The y axis from top to bottom ranks the variables’ importance, which is the mean of their absolute SHAP values.

Our model meets several criteria for effective early warning scores<sup>13</sup>—data is gathered primarily electronically, results can be reported in a timely manner, and results may trigger clinical actions to affect a proximate,

clinically meaningful outcome. Notably, even our clinical ND models rely on very little manual input, and we included only features that would be readily available on initial presentation. While information such as initial radiographic findings and bloodwork may improve predictive power, we intentionally excluded this information from analysis, since our objective was to identify methods that could also be applied in pre-hospital or austere environments. Our data was collected using wearable sensors already in use as part of standard clinical practice. We generated three rapid ND prediction models in order to improve versatility of application in cases where either physiologic or clinical data is unavailable.

The clinical features contributing to our combined ND prediction model included initial GCS and injury type. GCS is an important contributing variable in existing validated TBI prognostic models including CRASH and IMPACT scores<sup>17,18</sup>. Our finding that lower initial GCS is associated with a higher likelihood of early ND suggests that the relationship between initial GCS and outcome may be not only due to more severe primary injuries, but also to a higher likelihood of secondary ND. While associated with higher mortality, penetrating injury type is less commonly incorporated into TBI prognostic models<sup>46,47</sup>. There is a paucity of literature on the epidemiology and management of penetrating TBI, with some prognostic models excluding this injury type entirely, due to the assumption that it has a distinct pathophysiology. The fact that injury type contributes to our ND prediction model may indicate a need for more focused neuromonitoring in this population. Future work investigating physiologic differences between penetrating and non-penetrating TBI and the efficacy of earlier medical and surgical intervention in salvageable high-risk penetrating TBI patients may be warranted.

Interestingly, older age strongly contributed to increased risk for ND in the clinical model but only weakly contributed in the combined model. Older age consistently has been associated with worse outcomes after TBI<sup>48</sup>, and a recent single center study found a discordance between initial GCS, radiographic severity, and clinical outcome among elderly patients<sup>49</sup>, suggesting that a period of lucidity is more common among older patients who are then more likely to clinically deteriorate. Aging is associated with changes in heart rate variability<sup>50,51</sup>, so while numeric age was a lesser contributor to our combined model, physiologic aging was likely captured by other features.

Physiologic features that contributed to ND prediction included static admission VS, standard HRV features, and some more novel composite physiologic features. The strongest of these included the SD1 to SD2 ratio, the absolute power, and the low frequency power. The SD1 to SD2 ratio is a nonlinear dynamics feature derived from the Poincare plot—a scatterplot comparing each R-R interval to its previous interval<sup>34,52</sup>. It is thought to represent the unpredictability of the beat-to-beat time series and correlates with other HRV measures reflecting the balance between sympathetic and parasympathetic branches of the autonomic nervous system<sup>34</sup>. Our results suggest that a more predictable ECG time series with less physiologic variability is associated with higher likelihood of ND. The total power is a frequency domain feature which represents the variance (average of squared differences from the mean) of all R-R intervals<sup>29</sup>, with our results suggesting that high variance contributed to higher risk for ND in some patients. Low frequency power is another frequency-domain feature that is an accepted marker of sympathetic activity<sup>29,52</sup>, and our results suggest that increased early sympathetic activity was associated with higher likelihood of ND.

Describing all physiologic features contributing to this model is beyond the scope of this work and may become misleading, as the physiologic underpinnings of even the most standard measures remain somewhat opaque. Moreover, our machine learning data-driven model ultimately combines existing parameters into a new clinical-physiologic computational signal associated with ND.

Our physiologic variables included several calculated features derived from transfer learning to estimate the likelihood of clinical variables including GCS and ISS so that the decision support tools could run in a real time automated manner without losing important clinical information. The estimated GCS variables were developed from a large trauma cohort of approximately 28,000 patients<sup>35,36</sup>. This is the first time it has been applied to a TBI-specific population. As the feature's name implies, it was initially developed as a substitution for GCS in physiologic only models. However, the fact some of the estimated “physiologic GCS” features continue to contribute to the combined model where actual clinical GCS is also included highlights that these features likely capture something unique, beyond GCS. The GCS is a widely used but crude scoring system with a number of limitations including heterogeneity among patients with the same total score. For example, with each GCS total score from 4 to 14, more than one disorder of consciousness is not only possible but occurs with some frequency in the real world, as reported from TRACK-TBI data<sup>53</sup>. One explanation, then, is that the estimated GCS variables may relate more closely to the most frequent level of consciousness associated with each GCS score. Additionally, GCS inter-rater reliability even among neurologists is far from perfect<sup>54</sup>, so we may not expect a perfect correlation between clinical and physiologic GCS. Ongoing work will further evaluate the clinical relevance of these transfer learning features.

Univariable comparisons between patients with versus without ND demonstrated small but statistically significant differences in initial oxygen saturation and body temperature. These values cannot be used in isolation to predict ND. Through multivariable mathematical modeling we found that initial body temperature but not oxygen saturation contributed to ND prediction, as demonstrated in Fig. 5. This is consistent with a pre-hospital study demonstrating that initial body temperature in a large TBI cohort made a difference in predicting mortality<sup>55</sup>. Our findings support the need for machine learning tools to detect subtle but clinically relevant physiologic patterns that would be impossible for a human bedside clinician to recognize.

This is an exploratory proof of concept study with a number of limitations. As the initial phase of a multi-phase study, we aimed to generate robust models predicting neuro-specific decline in the setting of isolated but heterogeneous TBI. To that end, we excluded patients with major thoracic or abdominal trauma that might create systemic sources of ND. It is likely that ND prediction power may improve by stratifying patients based on injury severity, but here we sought a robust signal from TBI patients across the severity spectrum. We hope that our models may eventually serve as a rule-out decision support tool. In this study, we excluded patients



who were discharged before 48 h, who were deemed not to require ongoing inpatient monitoring by the clinical team. This resulted in loss of data from the mildest TBI cases, whom we believe unlikely to experience out-of-hospital ND based on conservative observation strategies. Other limitations of this study included retrospective data collection predisposing to reporting bias, which we attempted to minimize by selecting objective and well-documented criteria for ND. While this study included a large cohort that expands upon our previous pilot study work including a mostly mild cohort<sup>19</sup>, prospective and multicenter validation is needed. Without a testing validation cohort, our model may be overfit and performance over-estimated. Our models also do not account for differences in pre-hospital transport times, which were not reliably available for analysis. To address many of these shortcomings, we have begun a more inclusive prospective clinical validation trial (NCT: NCT05084352). Future efforts also may explore the use of ND physiologic prediction models in the pre-hospital setting<sup>16,35</sup>.

We reported model performance here as time-dependent areas under the receiver operating curve, but alternative clinically relevant performance metrics will be explored in the future. As an initial triage decision support tool, we prioritize sensitivity. However, specificity may be improved with the addition of portable and wearable neurosensors, such as reduced montage electroencephalography or pupillometry. However, widespread application of the physiologic models proposed here also will be limited by data infrastructure availability for recording and computing physiologic signals from bedside ECG and PPG monitors.

Further, these models do not represent prescriptive tools; they are capable of providing a risk score for ND but do not suggest a management strategy. This is in contrast to previous work by our group predicting the need for massive transfusion based on early physiologic data<sup>15,35</sup>. At this time, the ND prediction results simply provide early warning notification for near future care planning. Following model validation, simulation-based testing represents a next step; clinicians who work in the trauma resuscitation unit will be given a scenario, the bedside VS display, and a time-dependent ND risk score in order to determine whether and how these scores may affect management practices.

## Conclusions

Data-driven ML models based primarily on non-invasive physiologic monitoring during the “golden hour” of TBI resuscitation have potential to predict risk for subsequent ND and provide an opportunity to mitigate associated secondary injury and worse outcomes. These models may support time-sensitive decisions regarding triage, management, and resource utilization for acute TBI patients, with the goal of improving outcomes and reducing unnecessary costs. The model presented here is based on a large TBI cohort that expands upon previous smaller studies and, if validated, may emerge as an important clinical decision support tool.

## Data availability

This analysis is part of an ongoing USAF funded study and will be released to FITBIR after the prospective data collection is completed, which is anticipated by July 2024.

Received: 18 October 2022; Accepted: 13 December 2022

Published online: 09 January 2023

## References

- Brain Trauma Foundation. <https://www.braintrauma.org/faq>.
- Otten, E. J. & Dorlac, W. C. Managing traumatic brain injury: Translating military guidelines to the wilderness. *Wilderness Environ. Med.* **28**, S117–S123 (2017).
- Gurney, J. M. *et al.* The prehospital evaluation and care of moderate/severe TBI in the austere environment. *Mil. Med.* **185**, 148–153 (2020).
- Cowley, R. Resuscitation and stabilization of major multiple trauma patients in a trauma center environment. *Clin. Med.* **83**, 16–22 (1976).
- Stengel, D. *et al.* Point-of-care ultrasonography for diagnosing thoracoabdominal injuries in patients with blunt trauma. *Cochrane Database Syst Rev* **12**, CD012669 (2018).
- Morris, G. F., Juul, N., Marshall, S. B., Benedict, B. & Marshall, L. F. Neurological deterioration as a potential alternative endpoint in human clinical trials of experimental pharmacological agents for treatment of severe traumatic brain injuries. Executive Committee of the International Selfotel Trial. *Neurosurgery* **43**, 1369–1372 (1998) (**discussion 1372–1374**).
- Choudhry, O. J., Prestigiacomo, C. J., Gala, N., Slasky, S. & Sifri, Z. C. Delayed neurological deterioration after mild head injury: Cause, temporal course, and outcomes. *Neurosurgery* **73**, 753–760 (2013) (**discussion 760**).
- Majidi, S., Siddiq, F. & Qureshi, A. I. Prehospital neurologic deterioration is independent predictor of outcome in traumatic brain injury: Analysis from National Trauma Data Bank. *Am. J. Emerg. Med.* **31**, 1215–1219 (2013).
- Lingsma, H. F., Roozenbeek, B., Steyerberg, E. W., Murray, G. D. & Maas, A. I. R. Early prognosis in traumatic brain injury: From prophecies to predictions. *Lancet Neurol.* **9**, 543–554 (2010).
- Singer, K. E. *et al.* Efficacy of noninvasive technologies in triaging traumatic brain injury and correlating with intracranial pressure: A prospective study. *J. Surg. Res.* **262**, 27–37 (2021).
- Bazarian, J. J. *et al.* Serum GFAP and UCH-L1 for prediction of absence of intracranial injuries on head CT (ALERT-TBI): A multicentre observational study. *Lancet Neurol.* **17**, 782–789 (2018).
- Martín-Rodríguez, F. *et al.* Identification of serious adverse events in patients with traumatic brain injuries, from prehospital care to Intensive-Care Unit, using early warning scores. *Int. J. Environ. Res. Public Health* **17**, 1504 (2020).
- Kramer, A. A., Sebat, F. & Lissauer, M. A review of early warning systems for prompt detection of patients at risk for clinical decline. *J. Trauma Acute Care Surg.* **87**, S67–S73 (2019).
- Yang, S., Stansbury, L. G., Rock, P., Scalea, T. & Hu, P. F. Linking big data and prediction strategies: Tools, pitfalls, and lessons learned. *Crit. Care Med.* **47**, 840–848 (2019).
- Parimi, N. *et al.* Automated continuous vital signs predict use of uncrossed matched blood and massive transfusion following trauma. *J. Trauma Acute Care Surg.* **80**, 897–906 (2016).
- Zeineddin, A. *et al.* Prehospital continuous vital signs predict need for resuscitative endovascular balloon occlusion of the aorta and resuscitative thoracotomy prehospital continuous vital signs predict resuscitative endovascular balloon occlusion of the aorta. *J. Trauma Acute Care Surg.* **91**, 798–802 (2021).

17. Steyerberg, E. W. *et al.* Predicting outcome after traumatic brain injury: Development and international validation of prognostic scores based on admission characteristics. *PLoS Med.* **5**, e165 (2008) (**discussion e165**).
18. MRC CRASH Trial Collaborators *et al.* Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* **336**(7641), 425–429. <https://doi.org/10.1136/bmj.39461.643438.25> (2008).
19. Melinosky, C. *et al.* Continuous vital sign analysis to predict secondary neurological decline after traumatic brain injury. *Front. Neurol.* **9**, 761 (2018).
20. Baker, S. P., O'Neill, B., Haddon, W. & Long, W. B. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J. Trauma* **14**, 187–196 (1974).
21. Hu, P. *et al.* Reliable collection of real-time patient physiologic data from less reliable networks: A “Monitor of Monitors” system (MoMs). *J. Med. Syst.* **41**, 1–8 (2016).
22. Marshall, L. F. *et al.* The diagnosis of head injury requires a classification based on computed axial tomography. *J. Neurotrauma* **9**(Suppl 1), S287–292 (1992).
23. Maas, A. I. R., Hukkelhoven, C. W. P. M., Marshall, L. F. & Steyerberg, E. W. Prediction of outcome in traumatic brain injury with computed tomographic characteristics: A comparison between the computed tomographic classification and combinations of computed tomographic predictors. *Neurosurgery* **57**, 1173–1182 (2005) (**discussion 1173–1182**).
24. WMA - The World Medical Association-WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>.
25. Health Insurance Portability and Accountability Act of 1996 (HIPAA) | CDC. <https://www.cdc.gov/php/publications/topic/hipaa.html> (2022).
26. Fujimoto, K., Miura, M., Otsuka, T. & Kuratsu, J.-I. Sequential changes in Rotterdam CT scores related to outcomes for patients with traumatic brain injury who undergo decompressive craniectomy. *J Neurosurg* **124**, 1640–1645 (2016).
27. Yang, S., Hu, P. F. & Mackenzie, C. F. Trauma outcome prediction in the era of big data: From data collection to analytics. In *Signal Processing and Machine Learning for Biomedical Big Data* (eds Sejdíć, Ervin & Falk, Tiago H.) (CRC Press, 2018).
28. Garcia, D. Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput. Stat. Data Anal.* **54**, 1167–1178 (2010).
29. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* **93**, 1043–1065 (1996).
30. Krafty, R. T., Zhao, M., Buysse, D. J., Thayer, J. F. & Hall, M. Nonparametric spectral analysis of heart rate variability through penalized sum of squares. *Stat. Med.* **33**, 1383–1394 (2014).
31. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **15**, 70–73 (1967).
32. Laguna, P., Moody, G. B. & Mark, R. G. Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals. *IEEE Trans. Biomed. Eng.* **45**, 698–715 (1998).
33. Boardman, A., Schlindwein, F. S., Rocha, A. P. & Leite, A. A study on the optimum order of autoregressive models for heart rate variability. *Physiol. Meas.* **23**, 325–336 (2002).
34. Henriques, T. *et al.* Nonlinear methods most applied to heart-rate time series: A review. *Entropy (Base)* **22**, 309 (2020).
35. Yang, S. *et al.* Comparison of massive and emergency transfusion prediction scoring systems after trauma with a new bleeding risk index score applied in-flight. *J. Trauma Acute Care Surg.* **90**, 268–273 (2021).
36. Yang, S. *et al.* Utilizing ultra-early continuous physiologic data to develop automated measures of clinical severity in a traumatic brain injury population. Neurocritical Care Society 2022 Annual Meeting. *Neurocrit Care* **37** (Suppl 3), 329–717 (2022).
37. Pan, S. J. & Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2010).
38. Ridgeway, G. The state of boosting. in *Computing Science and Statistics* 172–181 (1999).
39. Moncada-Torres, A., van Maaren, M. C., Hendriks, M. P., Siesling, S. & Geleijnse, G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Sci Rep* **11**, 6968 (2021).
40. Peng, H., Long, F. & Ding, C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1226–1238 (2005).
41. Lundberg, S. & Lee, S.-I. A Unified approach to interpreting model predictions. [arXiv:1705.07874 \[cs, stat\]](https://arxiv.org/abs/1705.07874) (2017).
42. Pencina, M. J. & D'Agostino, R. B. Overall C as a measure of discrimination in survival analysis: Model specific population value and confidence interval estimation. *Stat. Med.* **23**, 2109–2123 (2004).
43. Haibe-Kains, B., Desmedt, C., Sotiriou, C. & Bontempi, G. A comparative study of survival models for breast cancer prognostication based on microarray data: Does a single gene beat them all?. *Bioinformatics* **24**, 2200–2208 (2008).
44. Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp 785–794 (2016).
45. Nielson, J. L. *et al.* Statistical guidelines for handling missing data in traumatic brain injury clinical research. *J. Neurotrauma* **38**, 2530–2537 (2021).
46. Santiago, L. A., Oh, B. C., Dash, P. K., Holcomb, J. B. & Wade, C. E. A clinical comparison of penetrating and blunt traumatic brain injuries. *Brain Inj.* **26**, 107–125 (2012).
47. Loggini, A. *et al.* Management of civilians with penetrating brain injury: A systematic review. *J. Crit. Care* **56**, 159–166 (2020).
48. Hukkelhoven, C. W. P. M. *et al.* Patient age and outcome following severe traumatic brain injury: An analysis of 5600 patients. *J. Neurosurg.* **99**, 666–673 (2003).
49. Garza, N., Toussi, A., Wilson, M., Shahlaie, K. & Martin, R. The increasing age of TBI patients at a single level 1 trauma center and the discordance between GCS and CT Rotterdam scores in the elderly. *Front. Neurol.* **11**, 112 (2020).
50. De Meersman, R. E. & Stein, P. K. Vagal modulation and aging. *Biol. Psychol.* **74**, 165–173 (2007).
51. Buchman, T. G., Stein, P. K. & Goldstein, B. Heart rate variability in critical illness and critical care. *Curr. Opin. Crit. Care* **8**, 311–315 (2002).
52. Shaffer, F. & Ginsberg, J. P. An overview of heart rate variability metrics and norms. *Front. Public Health* **5**, 258 (2017).
53. Bodien, Y. G. *et al.* Diagnosing level of consciousness: The limits of the Glasgow coma scale total score. *J. Neurotrauma* **38**, 3295–3305 (2021).
54. Fischer, M. *et al.* Inter-rater reliability of the Full Outline of UnResponsiveness score and the Glasgow Coma Scale in critically ill patients: A prospective observational study. *Crit Care* **14**, R64 (2010).
55. Gaither, J. B. *et al.* Body temperature after EMS transport: Association with traumatic brain injury outcomes. *Prehosp. Emerg. Care* **21**, 575–582 (2017).

## Acknowledgements

We would like to thank Yasmin Alamdeen and Devon Beagan for their research assistance contributions to this work.

### Author contributions

Study conceptualization and supervision: N.B., P.H.; Study design and planning: N.B., P.H., S.Y., G.P., J.P., C.M., H.T.; Clinical and demographic database management and data analysis: S.M., R.F.; Acquisition and interpretation of clinical and radiographic data: N.B., G.P., J.P., S.M., R.F., H.T.; Acquisition and processing of physiologic data: P.H., S.Y., Y.K., C.L., S.M., R.F.; Physiologic data analysis/modeling: S.Y., H.C.; Critical review of results: J.P., N.B., P.H., S.Y., H.C., R.F.; Manuscript writing: J.P.; Manuscript revision: N.B.; G.P.; S.Y.; P.H.; R.F. Final approval of manuscript: All authors.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26318-4>.

**Correspondence** and requests for materials should be addressed to N.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023