Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Early Clinical Neuroworsening is Associated with a Multidimensional Cerebral Edema-Related Endophenotype after Traumatic Brain Injury

Luisa M Rojas Valencia
University of California San Francisco
Esmeralda Mendoza
University of California San Francisco
Hannah L Radabaugh
University of California San Francisco
Sanhita Nittala
University of California Berkeley
Tyler Henry
Oregon Health & Science University
Dennis Obat
University of California San Francisco
John H. Kanter
University of Pittsburgh
John K Yue
University of California San Francisco
Jeffrey Pollock
Oregon Health & Science University
Cydni N Williams
Oregon Health & Science University
Cillian E. Lynch
Oregon Health & Science University
Ramon Diaz Arrastia
University of Pennsylvania
Susan Rowell
The University of Chicago
Adam R. Ferguson
University of California San Francisco; San Francisco Veterans Affairs Healthcare System
H.E. Hinson
h.e.hinson@ucsf.edu

UCSF: University of California San Francisco https://orcid.org/0000-0003-1598-9782

Research Article

Keywords: blood-based biomarkers, diagnosis, neuroworsening, neurological deterioration, traumatic brain injury

Posted Date: November 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11245359/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background

Early neurological deterioration after traumatic brain injury (TBI) is associated with poor outcome, yet identifying patients at risk for neuroworsening during the first 48 hours after injury can be challenging. Our goal was to identify patients that develop clinical or radiographic neuroworsening in the first 48 hours to determine factors that are associated with neuroworsening.

Methods

We analyzed patients from the Predictors of Low-risk Phenotypes after TBI Incorporating Proteomic Biomarker Signatures (PROTIPS) cohort of patients sustaining TBI (Glasgow Coma Scale (GCS) 3–12) confirmed by presence of hemorrhage on head computed tomography (CT) scan. Clinical neuroworsening was defined as a sustained decrease of ≥ 2 points in GCS in two consecutive measurements during the first 48 hours after admission. Demographic, clinical, radiographic, and biomarker injury characteristics were compared between the groups. A nonlinear principal component analysis (NL-PCA) was performed to identify endophenotypes present in the PROTIPS cohort. Neuroworsening groups were compared to determine if a specific endophenotype was associated with neuroworsening. Similar analyses were performed using one point increase in Rotterdam score as an endpoint to determine radiographic neuroworsening.

Results

Among the 123 patients, 40% experienced clinical neuroworsening. Patients in the clinical neuroworsening group had lower admission GCS and greater injury severity than patients without neuroworsening. NL-PCA identified four endophenotypes (neurovascular-endothelial injury, cerebral edema, pauci-inflammatory, and pro-inflammatory state), of which the cerebral edema endophenotype was significantly associated with clinical neuroworsening ($p = 0.002$, Cohen's $d = 0.6$) and demonstrated modest discriminatory performance (AUC = 0.67). When radiographic neuroworsening was defined using Rotterdam scores, NL-PCA identified four endophenotypes: vascular injury and edema, pro-inflammatory, metabolic/endothelial activation, and angiogenic. No significant differences were found in radiographic neuroworsening across the PCs.

Conclusions

Early clinical neuroworsening in GCS 3–12, CT-positive TBI was associated with a multidimensional cerebral edema-related endophenotype over the first 48 hours post-injury, rather than any individual feature alone.

Introduction

Traumatic brain injury (TBI) is a leading cause of death and long-term disability worldwide [1], with a lifetime prevalence of TBI with loss of consciousness of approximately 18.2% in the US [2]. In the acute clinical setting, clinicians must rapidly make decisions within the first hours after injury regarding triage, need for imaging, and transfer to a higher level of care. These decisions have the potential to dramatically alter the patient's eventual outcome. While most patients recover with supportive care, a significant fraction do not and require intensive surgical and critical care management. Therefore, early identification of patients at risk of neurological deterioration is essential [3].

The mechanisms of neurological deterioration or "neuroworsening" arise from secondary brain injury after TBI and include hypoxia, neuroinflammation, cerebral edema, ischemia, excitotoxicity, and others [4]. There are no universally accepted criteria for defining or diagnosing neuroworsening [5]. Neuroworsening has been defined by some authors as the decrease of 2 or more points in Glasgow Coma Scale (GCS) motor score [6], decrease of 2 or more points on the total GCS score [5], or progression in radiographic findings like subarachnoid hemorrhage (SAH), intraparenchymal hemorrhage (IPH), subdural hemorrhage (SDH) and epidural hematoma (EDH) [7, 8], or some combination of clinical and radiographic criteria [9]. Early neuroworsening (within 48 hours) has been shown to be associated with longer-term outcome, emphasizing the clinical importance of identifying patients likely to deteriorate, so that they can be monitored more closely and managed proactively [10]. However, there is no predictive model validated to guide clinicians in the detection of acute neuroworsening. Clinicians would benefit from a method to screen and detect patients at risk of neuroworsening before it becomes clinically apparent.

To address this gap, we first evaluated two definitions of neuroworsening in a prospective, observational cohort of moderate/severe participants after TBI. The first definition was clinical neuroworsening: decrease in total GCS score by $\geq$ 2 points sustained over 2 measurements. The second definition was radiographic neuroworsening: an increase in the Rotterdam score by 1 point on successive head computed tomography (CT) scans during the first 48 hours. Then, with these clinical and radiographic neuroworsening definitions, we sought to describe the dominant biological drivers of each type of neuroworsening using a comprehensive set of patient data available upon admission (e.g. biomarker levels, CT findings, clinical assessments, etc.). Nonlinear principal component analysis (NL-PCA) was applied as an unsupervised dimensionality-reduction approach to identify major patterns of covariation within baseline clinical and biomarker data, representing endophenotypes. Subsequent PC loading analyses were used to interpret the biological features reflected by these endophenotypes. We then assessed the sensitivity of baseline endophenotypes for predicting clinical or radiographical neuroworsening.

Methods

Participants and data source:

Participants were enrolled in PRedictors Of low-risk phenotypes after Traumatic brain injury Incorporating Proteomic biomarker Signatures (PROTIPS). PROTIPS was a prospective, single-center (Oregon Health and Science University), observational cohort study that included adults and older teenagers (age>14) with GCS 3-12 on presentation due to TBI with intracranial hemorrhage present on the initial post-injury Head CT. Exclusion criteria were a GCS score of 3 without a reactive pupil, unknown time of injury, prisoners, pregnant individuals, or no hemorrhage on first head CT after injury. The participant's legally authorized representative provided consent for participation. If/when the participant regained the ability to provide informed consent, the participants were approached for consent. GCS scores were recorded hourly over the first 48 hours after admission. CT scan features were interpreted by neuroradiologists. Separately, research staffscored scans for injury severity using the Rotterdam scoring system[11], blind to the radiographic neuroworsening data analysis. Research coordinators extracted demographic data, injury characteristics, and hospital course information from the electronic medical record. This study was designed and powered for an *a priori* analysis of progression of intracranial hemorrhage as the primary endpoint with neuroworsening specified as an important secondary endpoint.

In the current report, we include effect sizes and observed power for each definition of neuroworsening to help inform future studies. It was estimated that 35% of participants would exhibit neuroworsening by the clinical definition. In the current analyses, 123 patients were assessed for clinical neuroworsening and 119 for radiographic neuroworsening due to data characteristics and missingness that limited the ability to categorize patients by the assigned outcome definitions (e.g., GCS score consistent at 3 across the 48-hour window or missing Rotterdam data, respectively) (Figure 1).

Biomarker analysis:

Patients were screened up to 3 hours from injury, and blood was drawn at the time of admission, centrifuged, and the plasma aliquoted and immediately stored at -80°C until batch processing. Analyte concentrations were measured using a combination of single- and multi-plex ultrasensitive electrochemiluminescence (ECL) kits run on the Mesoscale Discovery (MSD) QuickPlex SQ-120 platform (MesoScale Diagnostics LLC, Rockville, Maryland). See supplemental material for additional information, but in brief, levels of Tau (total), Glial Fibrillary Acidic protein (GFAP), and Neurofilament Light (NfL) were quantitated using the S-Plex Neurology Panel 1. Concentrations of apolipoprotein A (APOA) were measured via U-PLEX. Concentrations of C-Reactive Protein (CRP), and von Willebrand Factor (vWF) were multiplexed in U-PLEX format in combination with those of Intracellular Adhesion Molecule (ICAM-1), Serum Amyloid A (SAA), and Vascular Cell Adhesion Protein (VCAM-1). The R-PLEX analytes Angiopoietin-1 (Ang-1), Angiopoietin-2 (Ang-2), Thrombomodulin, Syndecan-1, and Thrombospondin-2 (TSP-2) were multiplexed in combination with the U-PLEX analytes Interleukin-6 (IL-6), and tumor necrosis factor alpha (TNF-α). All samples were run in duplicate, alongside an eight-point calibration curve in duplicate to which the analyte concentrations for every individual plate were fit. The average value of the duplicates returned for each analyte and experimental sample was used as the final value

for each biomarker. Values are reported as picograms per milliliter (pg/mL). Blood biomarker concentration values were log transformed for analysis (see Supplementary Text S1).

Neuroworsening classification:

The primary outcome of neuroworsening was defined two ways. First, clinical neuroworsening was defined as a decrease of 2 or more points in the GCS score, sustained for at least two consecutive measurements over the first 48 hours of hospitalization, similar to previous reports [5]. We conducted a missingness analysis of the total GCS score measured hourly from time of admission to 48 hours to determine patterns of missingness and inform downstream data handling and imputation where appropriate (using the multivariate imputation by chained equations (MICE) package; R 4.5.1 (R Core Team, 2021). Then, the patients were categorized by their total GCS scores over time into either the neuroworsening group (decrease of ≥ 2 sustained over two recordings) or no-neuroworsening group.

In a completely independent analysis, radiographic neuroworsening was defined as a one-point increase in Rotterdam score on subsequent CT scan. The 119 participant trajectories were divided into unchanging (unchanging score over the 4 time periods – baseline, 2-6 hours, 7-24 hours, and 25-48 hours) or changing for those whose Rotterdam scores have any variation over time. Neuroworsening was adjudicated by an expert neurologist (HEH) as an increase by 1 point or more in the Rotterdam scores during the timepoints previously described. The previous classification generated 3 groups of patients: unchanging (no change in Rotterdam score), changing but not worsening (decreasing Rotterdam score over time), and changing and worsening (increasing Rotterdam score over time) (Supplementary Figure 1).

Descriptive Statistics:

After neuroworsening groups were defined (as above), descriptive statistics were used to report continuous variable means and standard deviations ($\pm$SD) or medians and interquartile ranges [Q1, Q3] and categorical variable proportions (%) for each Neuroworsening group (Tables 1 and 2). For two group comparisons (clinical neuroworsening), continuous variables were compared using Wilcoxon rank-sum, and categorical variables were compared using Pearson's chi-squared or Fisher's exact test. For three group comparisons (radiographic neuroworsening), continuous variables were compared using Kruskal-Wallis rank-sum test and categorical variables were compared using Fisher's exact test. Given that the two definitions of neuroworsening differ by which variable was used to determine the outcome categorization (i.e., total GCS score vs Rotterdam score), variables included as input into each of the independent analyses differ to avoid circularity and outcome leakage between input variables and specific output. Variables included for each independent NL-PCA can be observed in Figures 2 and 3.

Principal Component Analysis (NL-PCA):

Principal component analyses (NL-PCA) using the syndRomics package (R version 4.5.1) [12] were performed separately on each set of variables (Figures 2-3) to reduce dimensionality, circumvent

multicollinearity, and identify dominant multivariate covariance patterns within each set of admission variables for subsequent testing of their associations with the relevant neuroworsening outcome. Scree plots from the eigenvalues and variance were plotted to select the number of principal components (PCs) to advance through the analysis stages. PCs were selected using both the elbow-method which observes the asymptote in the scree plot and through bootstrapped permutation testing to identify the PCs in which variance accounted for (VAF) was greater than chance [12]. Among the selected PCs, a loading threshold of |0.45| was used to label the PCs by descriptive endophenotype (i.e., assign a clinically interpretable name informed by the dominant variable loadings). Each participant was assigned a PC score for each endophenotype label.

Association of Endophenotypes and Clinical Variables with Neuroworsening Outcomes:

For the clinical neuroworsening (i.e., GCS score) NL-PCA, a two-sample t-test was used to identify significant ($p < 0.05$) group differences across endophenotypes (i.e., PC scores) for each selected PC. For the radiographic neuroworsening (i.e., Rotterdam score) NL-PCA, an Analysis of Variance (ANOVA) was used to identify significant group differences in endophenotype (i.e., PC scores) across the three Rotterdam trajectory groups (unchanging, changing but not worsening, and changing and worsening). A threshold of $p<0.05$ was considered statistically significant. Data were analyzed (using libraries stats, pROC, and ROCR) and visualized as heatmaps, bar maps, and syndromic plots (libraries, ggplot2, Gifi, progress, and syndRomics) in R version 4.5.1.

Following NL-PCA, PCs that demonstrated significant differences between participants with and without the binary definition used for clinical neuroworsening (GCS Score) were examined further to determine whether their association with the defined outcome was attributable to individual variables or reflected a broader multivariable pattern. For each significant PC, variables with absolute loadings ≥0.45 were identified as the dominant contributors to that component and entered simultaneously into a multivariable logistic regression model with clinical neuroworsening as the binary outcome. This analysis was used to evaluate whether individual variables within the NL-PCA-defined endophenotype were independently associated with neuroworsening after accounting for the other dominant contributors to the component. R-squared, significance, p-values, 95% confidence intervals, and coefficients for the logistic regression model were determined for each of the two neuroworsening definitions. Receiver operating characteristic (ROC) curves were generated to characterize the ability of neuroworsening-associated PC scores and subsequent multivariable logistic regression models to discriminate between participants with and without clinical neuroworsening. Discrimination was summarized using the area under the ROC curve (AUC).

Results

Participants

In the PROTIPS study, 241 patients were screened for eligibility. Of these, 115 were excluded based on predefined criteria: absence of hemorrhage on CT (n=79), declined consent (n=8), survival <24 hours

(n=14), unable to obtain a blood sample within 3 hours of trauma (n=9), and other exclusion criteria (n=5). A total of 126 subjects were enrolled. Three patients were excluded from the clinical neuroworsening, GCS analysis because their GCS score remained at 3 throughout the entire 48-hours of admission, and thus could not deteriorate according to our definition (n=123). Seven patients were excluded from the analysis of the radiographic neuroworsening using Rotterdam scores due to the fact that only a single Rotterdam score was available for analysis, thus deterioration could not be assessed (n=119). A detailed flow of patient inclusion and exclusion is provided in (Figure 1).

Demographics and Injury Characteristics of Clinical Neuroworsening

Among the 123 patients who were included in the analysis, 49 (40%) experienced clinical neuroworsening. There were no significant differences in age, sex, race, or ethnicity between clinical neuroworsening and non-neuroworsening patients (Table 1). Subjects in the clinical neuroworsening group (n=49) had lower admission Glasgow Coma Scale (GCS) scores (7 [3, 10] vs. 8[7, 11], p=0.009), higher total body Injury Severity Score (ISS) (26 [18, 33] vs. 21 [14, 29], p=0.02), and more frequent intracranial monitoring (57% vs. 27%, $p < 0.001$) compared to the no-neuroworsening group. In univariate analysis, the clinical neuroworsening group had significantly higher Rotterdam scores (3 [3, 4] vs. 3 [3, 3], $p = 0.002$). Among circulating biomarkers, Ang-1 and Syndecan-1 were significantly higher in patients with clinical neuroworsening (p=0.02 and p=0.04). Participants with clinical neuroworsening also were more likely to die during their hospitalization and have a worse 6-month outcome as measured by the 6-month Glasgow Outcome Scale–Extended (GOSE) (3 [1, 4] vs. 4 [3, 7], $p = 0.004$) (Table 1).

Principal component analysis identified 4 distinct endophenotypes

The NL-PCA analysis for the clinical neuroworsening outcome included demographic, clinical, radiographic, and blood biomarker data (Table 1). Four PCs were retained for analysis of association with clinical neuroworsening based on the Scree criterion and permutation testing (Supplementary Figure 2). The first 4 PCs explained 63% of the total variance, with PC1 representing 33.5% of the variance (Figure 2A). PC1 was labeled as “neurovascular–endothelial injury” (Figure 2A). Variables with high loadings in PC1 (>|0.45|) included vWF, TSP-2, Thrombomodulin, TNF-α, APOA, Tau, Ang-1, Ang-2, Syndecan-1, VCAM-1 and ICAM-1. PC2 (12.8% VAF) was labeled “cerebral edema” given the high loadings of midline shift, compressed basal cisterns, high Head Abbreviated Injury Severity Score (HAIS), and elevated admission International Normalized Ratio (INR), GFAP and NfL (Figure 2B). PC3 (8.8% VAF) was assigned “pauci-inflammatory” as higher hemoglobin and higher platelets loaded in the opposite direction of younger age, lower CRP, INR, and SAA (Figure 2C). Finally, PC4 (8.0% VAF) was assigned “pro-inflammatory” given elevated CRP, elevated platelets, and elevated SAA loading together (Figure 2D).

Cerebral edema endophenotype is associated with clinical neuroworsening

Associations were examined between each PC and clinical neuroworsening. Among the four PC endophenotypes, the cerebral edema endophenotype (PC2) was the only one that was significantly

associated with clinical neuroworsening (p = 0.002), with a moderate effect size (Cohen's d = 0.60) (Figure 2E**)**. PC2 was primarily driven by GFAP, NfL, basal cisterns, HAIS, IL-6, midline shift, and INR at admission. The cerebral edema endophenotype demonstrated modest discrimination for clinical neuroworsening, with an area under the curve (AUC) of 0.67 (Figure 2F). This indicates that patients who experienced neuroworsening generally had higher PC2 scores than those who did not.

To further evaluate the predictive value of the cerebral edema endophenotype, a supervised logistic regression model incorporating the individual variables with absolute loading values ≥ 0.45 on PC2 as input and clinical neuroworsening as output was constructed. Although none of the individual predictors were independently associated with neuroworsening, the combined model demonstrated modest discriminatory ability (AUC=0.72) (72%). Together, these findings are consistent with the association between the cerebral edema endophenotype and clinical neuroworsening reflecting a multivariable pattern rather than a single dominant marker. (Figure 2F). No significant differences were observed for neurovascular-endothelial injury (PC1), pauci-inflammatory (PC3), and pro-inflammatory state (PC4) (p = 0.14, 0.48, and 0.23, respectively).

The changing but not worsening radiographic group appeared to be the most severely injured group

There were no differences in demographics or admission total GCS between the three radiographic neuroworsening groups (Table 2). However, those participants with changing but not worsening Rotterdam scores had higher HAIS scores, lower GCS motor scores, and the highest Rotterdam scores signaling more severe injury on presentation, suggesting a ceiling effect for detection of neuroworsening. Congruently, the changing but not worsening group also had higher levels of the blood-based biomarkers GFAP and NfL. The in-hospital mortality of this group was the highest of the three, 50% vs. 25% in the changing and worsening group, and vs. 12% in the unchanging group (p= 0.006) (Table 2).

Principal component analysis identified 4 distinct radiographic endophenotypes

The radiographic neuroworsening NL-PCA analysis included demographic, clinical, total GCS at admission, and blood biomarker variables, importantly eliminating radiographic variables that would introduce circularity. Four PCs, explaining 73.8% of the total variance in the dataset, were kept for radiographic neuroworsening analyses based on the Scree criterion and permutation-based testing (Supplementary Figure 3). PC1 (35.9% VAF) was assigned a "vascular injury and edema" endophenotype, with high ICAM-1, VCAM-1, Ang-1, Ang-2, Syndecan-1, TSP-2, Thrombomodulin, TNF-α, APOA, Tau, and GFAP as dominant loaders (Figure 3A). PC2 (18.2% VAF) was assigned a "pro-inflammatory" endophenotype, with high CRP, platelets, SAA, NfL, and fibrinogen, loading in the opposite direction of hemoglobin (Figure 3B). PC3 (11.9% VAF) represents a "metabolic/endothelial activation" endophenotype, driven by low glucose, high vWF and high IL-6(Figure 3C). PC4 (7.8% VAF) was assigned an "angiogenic" endophenotype, with high Ang-1, low GCS score, younger age, and high head abbreviated injury score (HAIS) (Figure 3D). An ANOVA test comparing the three radiographic neuroworsening groups across the four PC endophenotypes showed no significant differences for vascular injury and

edema (PC1), pro-inflammatory (PC2), metabolic/endothelial activation (PC3), or angiogenic (PC4) (p = 0.88, 0.77, 0.76 and 0.53, respectively).

Discussion

In this study, we examined two complementary ways of conceptualizing neuroworsening: clinical deterioration, defined by worsening GCS scores (clinical neuroworsening) and radiographic deterioration, defined by worsening Rotterdam scores (radiographic neuroworsening). Using these definitions and the recently proposed Clinical, Biomarkers, Imaging and Modifiers (CBI-M)[13] conceptual framework, we applied NL-PCA on baseline features to characterize patterns of covariation across clinical variables, blood biomarkers, imaging features, demographics, and other potential disease modifiers. Clinical neuroworsening was associated with higher levels of Syndecan-1 and Ang-1 and worse 6-month functional outcome (lower GOSE scores). NL-PCA identified a cerebral edema endophenotype (PC2), driven by elevated GFAP, NfL, IL-6, crowded basal cisterns, high HAIS, worsening midline shift, and others which was associated with clinical neuroworsening over the first 48 hours. While no individual predictor reached statistical significance in the multivariable regression, the combined model demonstrated moderate discrimination (AUC 0.72). This pattern is consistent with clinical neuroworsening reflecting a multivariable injury endophenotype rather than an association dominated by a single marker.

Recent prior work using GCS-based neuroworsening definitions have shown similar associations with severity and outcome. Yue and colleagues defined neuroworsening as a decline in motor GCS at emergency department (ED) disposition, and found that neuroworsening is associated with subdural, subarachnoid, and intraventricular hemorrhage[10]. Neuroworsening also correlated with injury severity, need for neurosurgical intervention, and predicted unfavorable 3- and 6-month outcomes [10]. Prior work also identified higher GFAP concentrations among patients who worsened [14]. The present analysis extends those observations by integrating clinical, biomarker, imaging, and modifier features within a single framework. Rather than identifying GFAP or another isolated marker as the dominant signal, neuroworsening clustered with a coordinated edema-related pattern, suggesting a biologically coherent, multimodal representation of early deterioration.

The temporal relationship is biologically plausible given the evolution of cerebral edema during the early post-injury period [15], but the edema-related endophenotype should not be interpreted as evidence that edema is the sole cause of declining GCS. Its composition instead suggests convergence of structural mass effect, astroglial and axonal injury, inflammation, and systemic physiologic disturbance. The distinction is clinically important because neuroworsening in the intensive care unit (ICU) may also reflect hematoma expansion, seizure, hypercapnia, and sedation or analgesia, and other competing causes. A prospectively validated multimodal signature capable of distinguishing edema-related deterioration from these alternatives could support earlier reassessment and targeted management, including patients who may benefit from anti-cerebral edema therapies.

Discriminating neuroworsening by our definition of radiographic deterioration highlighted three groups of participants with different levels of injury severity and resultant outcomes, with the changing but not worsening group being the most severely injured and having the poorest outcomes. This might be explained by the observation that these patients had the highest Rotterdam scores of the three groups, and perhaps further radiographic deterioration cannot be detected at the upper limit of the score (ceiling effect). Missingness may also have influenced these results, as not all participants had serial head CTs and thus did not meet inclusion criteria. Some participants might have been too ill to travel to scans or too clinically stable to warrant repeat imaging. We did not identify a discrete endophenotype associated with the radiographic definition of neuroworsening. This might be because group differences were driven by overall severity of injury rather than a single pathophysiology. It is noteworthy that we excluded the Rotterdam radiological features used to define Radiographical Neuroworsening from the endophenotype NL-PCA to prevent circular analysis. This raises the possibility that Rotterdam features reflect orthogonal variance from the imaging endophenotype, providing distinct clinical value. Other investigators have used radiographic variables to determine neuroworsening [7] and have incorporated CT features for outcome prediction using NL-PCA [8] but not the specific evaluation of neuroworsening that we employed, limiting cross-study comparisons. Composite scores may not be the ideal endpoint for capturing radiographic neuroworsening. Instead, future studies might select other outcome variables for radiographic neuroworsening, such as increase in hematoma or edema volume on imaging, which might not suffer the same ceiling effect of ordinal scales.

Interestingly, across both clinical and radiographic models, common patterns emerged in the two NL-PCAs, including cerebral edema and inflammation. These endophenotypes may reflect mechanisms of secondary brain injury with differing temporal and spatial dynamics. Helping clinicians to identify these endophenotypes early in a patient's presentation may allow clinicians to be more proactive implementing intracranial monitoring of intracranial pressure or brain tissue oxygenation or deciding to perform decompressive surgery. Identification of these mechanistic endophenotypes also has important implications for clinical research developing therapies targeting these mechanisms. However, these associations must be externally validated prior to clinical implementation.

There are several important limitations to consider in the interpretation of this work. First, this study is limited by a single center design with a modest number of participants. External validation studies should include larger, multicenter cohorts. Radiographic neuroworsening was only observed in a very small number of participants, which raises the question of whether a composite score such as the Rotterdam score is insensitive to important clinical change. It is likely that other radiographic measures would be more sensitive to neuroworsening and may be better endpoints to reflect radiographic worsening. Alternatively, this study could be underpowered to detect radiographic worsening, or our sample did not reflect the best population to sample, but this seems less likely given the dynamic results we observed in the clinical neuroworsening arena. Another limitation is the exclusion of high dimensional electrophysiologic data such as invasive neuromonitoring data (e.g., intracranial pressure). The PROTIPS cohort represents data from a center with a less aggressive approach to neuromonitoring —only 39% of the cohort (48/123) had intracranial monitors. While this might not reflect the practice in

all academic centers, it improves the generalizability of our results to sites and specific patients where invasive neuromonitoring is not accessible.

Conclusion

Early clinical neuroworsening after CT-positive GCS 3–12 TBI was associated with a multidimensional cerebral edema-related endophenotype integrating clinical, imaging, and biomarker features within the CBI-M Framework. No individual feature independently distinguished patients who worsened, supporting the interpretation that the observed association reflects a distributed multivariate pattern rather than a single dominant marker. These findings warrant validation in larger, multicenter TBI cohorts prior to evaluation for clinical implementation and risk assessment.

Declarations

Author's contributions: All of the authors made substantial contributions to the conception or design of the study. Luisa M. Rojas Valencia, and Esmeralda Mendoza, did data analysis, interpretation, drafting and revision of the manuscript. Hannah L. Radabaugh, Sanhita Nittala, Tyler Henry, and Dennis Obat, contributed to data acquisition, analysis, and interpretation. John K. Yue and Jeffrey Pollock contributed to interpretation of data. John Kanter, Cydni Williams, and Cillian E. Lynch, contribute to data acquisition. Ramon Diaz-Arrastia, Susan Rowell, Adam R. Ferguson, contributed to data interpretation, drafting or revising the manuscript. H. E. Hinson contribute to data acquisition, analysis, interpretation, drafting and revising the manuscript (all authors).

This manuscript has not been published elsewhere and is not under consideration by another journal.

Confirmation of adherence to ethical guidelines: This study was IRB approved and adhered to the local Oregon Health and Science University IRB committee guidelines.

Financial Disclosure: John K. Yue received funding from the University of California, San Francisco Weill Institute Clinician-Scientist Award (#7032139) and the Congress of Neurological Surgeons Foundation Sophysa Young Investigator Award (#P0594482). Ramon Diaz-Arrastia received funding from U01 NS131740, U01 NS113140, R01 NS125408, and HT945-23-1-1039. Cydni Williams received funding from the National Heart Lung and Blood Institute (R01HL174653, K23HL150229). Adam R. Ferguson received funding from NIH/NINDS (U24NS122732; UH3NS106899; R01NS142318), Veterans Affairs (I01BX005871; I01RX002787; I01RX002245; I50BX005878), the Craig H. Neilsen Foundation, and the Wings for Life Foundation. H. E. Hinson received funding from HT9425-25-1-0336, HT9425-25-1-0550, and W81XWH-15-9-0001. Research reported in this publication was supported by the National Institute Of Neurological Disorders And Stroke of the National Institutes of Health under Award Number K23NS110828 (HEH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of interest: All the other authors declare no conflicts of interest.

Artificial intelligence statement: During the preparation of this manuscript, the author(s) used UCSF ChatGPT Enterprise in order to improve readability and language style. After using this tool the author(s) reviewed and edited heavily the content as needed and take full responsibility for the content of the published article.

Confirmation of the use of a reporting checklist: The STROBE checklist cohort study is attached.

References

1. Dewan MC, Rattani A, Gupta S, Baticulon RE, Hung Y-C, Punchak M, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg. 2019;130:1080–97. https://doi.org/10.3171/2017.10.JNS17352.
2. Karamian A, Lucke-Wold B, Seifi A. Prevalence of Traumatic Brain Injury in the General Adult Population of the USA: A Meta-Analysis. Neuroepidemiology. 2024;59:558–67. https://doi.org/10.1159/000540676.
3. Podell J, Yang S, Miller S, Felix R, Tripathi H, Parikh G, et al. Rapid prediction of secondary neurologic decline after traumatic brain injury: a data analytic approach. Sci Rep. 2023;13:403. https://doi.org/10.1038/s41598-022-26318-4.
4. Hinson HE, Rowell S, Schreiber M. Clinical evidence of inflammation driving secondary brain injury: A systematic review. J Trauma Acute Care Surg. 2015;78:184–91. https://doi.org/10.1097/TA.0000000000000468.
5. Vásquez-García S, Martínez-Palacios K, Hawryluk GWJ, Rabinstein A, Godoy DA. Neuroworsening in Traumatic Brain Injury: A Scoping Review of Definition, Prevalence, and Outcome. Neurocrit Care. 2025;43:424–36. https://doi.org/10.1007/s12028-025-02234-z.
6. Godoy DA, Fossi F, Robba C. Neuroworsening in Moderate Traumatic Brain Injury. Neurol Clin. 2025;43:51–63. https://doi.org/10.1016/j.ncl.2024.08.002.
7. Dalle Ore CL, Rennert RC, Schupper AJ, Gabel BC, Gonda D, Peterson B, et al. The identification of a subgroup of children with traumatic subarachnoid hemorrhage at low risk of neuroworsening. J Neurosurgery: Pediatr. 2018;22:559–66. https://doi.org/10.3171/2018.5.PEDS18140.
8. Yuh EL, Cooper SR, Ferguson AR, Manley GT. Quantitative CT Improves Outcome Prediction in Acute Traumatic Brain Injury. J Neurotrauma. 2012;29:735–46. https://doi.org/10.1089/neu.2011.2008.
9. Maas AIR, Murray G, Henney H, Kassem N, Legrand V, Mangelus M, et al. Efficacy and safety of dexanabinol in severe traumatic brain injury: results of a phase III randomised, placebo-controlled, clinical trial. Lancet Neurol. 2006;5:38–45. https://doi.org/10.1016/S1474-4422(05)70253-2.
10. Yue JK, Krishnan N, Kanter JH, Deng H, Okonkwo DO, Puccio AM, et al. Neuroworsening in the Emergency Department Is a Predictor of Traumatic Brain Injury Intervention and Outcome: A TRACK-TBI Pilot Study. J Clin Med. 2023;12:2024. https://doi.org/10.3390/jcm12052024.

11. Maas AIR, Hukkelhoven CWPM, Marshall LF, Steyerberg EW. 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. 2005;57:1173–82. discussion 1173–1182.
12. Torres-Espín A, Chou A, Huie JR, Kyritsis N, Upadhyayula PS, Ferguson AR. Reproducible analysis of disease space via principal components using the novel R package syndRomics. eLife. 2021;10:e61812. https://doi.org/10.7554/eLife.61812.
13. Manley GT, Dams-O'Connor K, Awwad H, Doperalski A, Umoh N, Maas A, et al. Marking a New Age in Characterization of Acute Traumatic Brain Injury: The National Institute of Neurological Disorders and Stroke Traumatic Brain Injury Classification and Nomenclature Initiative. J Neurotrauma. 2025;42:1021–2022. https://doi.org/10.1089/neu.2025.0134.
14. Yue JK, Fu AY, Elguindy MM, Van Essen TA, Eagle SR, Caldwell DJ, et al. Neuroworsening from a normal Glasgow Coma Scale Motor Score in the emergency department is an early predictor of neurosurgical intervention, hospital outcomes, and longitudinal disability in traumatic brain injury: A TRACK-TBI Study. Injury. 2026;113089. https://doi.org/10.1016/j.injury.2026.113089.
15. Jha RM, Kochanek PM, Simard JM. Pathophysiology and treatment of cerebral edema in traumatic brain injury. Neuropharmacology. 2019;145:230–46. https://doi.org/10.1016/j.neuropharm.2018.08.004.

Tables

Table 1. Demographic, clinical, and biomarker information by GCS neuroworsening group. Comparison of patients with and without neuroworsening as defined by drop in GCS by 2 pts or more in first 48 hours.

Characteristic	no-Neuroworsening N = 74[1]	Neuroworsening N = 49[1]	p-value[2]	Effect size[3]
Age at admission	48 (22)	44 (20)	0.3	0.182
Sex			0.3	0.177
Male	50 (68%)	37 (76%)		
Female	24 (32%)	12 (24%)		
Race			0.5	0.470
Asian	2 (2.7%)	3 (6.1%)		
Black	5 (6.8%)	6 (12%)		
Native American or Native Alaskan	1 (1.4%)	1 (2.0%)		
Pacific Islander or Native Hawaiian	2 (2.7%)	0 (0%)		
White	58 (78%)	38 (78%)		
Other	3 (4.1%)	1 (2.0%)		
Decline to answer	3 (4.1%)	0 (0%)		
Ethnicity			0.7	0.238
Hispanic or Latinx	5 (6.8%)	3 (6.1%)		
Non-hispanic, non-Latinx	67 (91%)	46 (94%)		
Decline to answer	2 (2.7%)	0 (0%)		
Admission GCS Score	8 [7, 11]	7 [3, 10]	0.009	0.497
GCS Motor	7 [7, 7]	7 [6, 7]	0.10	0.157
Injury Severity Score	21 [14, 29]	26 [18, 33]	0.020	0.457
Craniotomy			0.7	0.099
No	7 (33%)	8 (38%)		
Yes	14 (67%)	13 (62%)		
Intracranial Monitoring			<0.001	0.641
No	54 (73%)	21 (43%)		
Yes	20 (27%)	28 (57%)		

Characteristic	no-Neuroworsening N = 74[1]	Neuroworsening N = 49[1]	p-value[2]	Effect size[3]
Head Abbreviated Injury Score (HAIS)	4.0 [3.0, 4.0]	4.0 [4.0, 5.0]	0.007	0.591
Rotterdam	3.0 [3.0, 3.0]	3.0 [3.0, 4.0]	0.002	0.320
GFAP (pg/mL)	2,316.6 [632.1, 7,348.3]	3,827.8 [1,760.8, 11,300.0]	0.062	0.311
Ang-1 (pg/mL)	8,172.1 [5,247.7, 11,832.8]	10,545.7 [6,985.0, 16,192.1]	0.026	0.391
Ang-2 (pg/mL)	2,713.8 [2,035.2, 3,630.9]	2,524.2 [1,979.7, 3,041.8]	0.3	0.226
Syndecan-1 (pg/mL)	2,595.7 [2,139.4, 3,739.4]	3,119.2 [2,390.7, 4,279.2]	0.048	0.158
Serum Amyloid A (ng/mL)	2,076.7 [1,292.4, 4,569.5]	2,889.9 [1,312.3, 7,902.5]	0.3	0.233
ICAM-1 (pg/mL)	419,233.5 [334,523.6, 541,137.6]	411,022.8 [304,491.0, 523,479.1]	0.5	0.173
VCAM-1 (pg/mL)	906,133.9 [694,731.1, 1,112,177.5]	863,663.3 [673,301.1, 1,097,281.3]	0.6	0.065
vWF (ng/mL)	3,865.6 [2,826.1, 5,506.3]	3,274.9 [2,751.4, 4,929.3]	0.4	0.174
APOA (ng/mL)	1,980,489.6 [1,430,339.7, 2,471,132.2]	1,725,544.5 [1,341,029.5, 2,263,547.9]	0.14	0.156
IL-6 (pg/mL)	19.8 [12.4, 31.9]	26.4 [14.4, 42.9]	0.13	0.265
TNF-a (pg/mL)	8.5 [5.4, 14.3]	8.5 [6.2, 13.4]	0.8	0.165
CRP (ng/mL)	1,923.9 [618.0, 5,122.1]	1,727.5 [1,152.3, 4,062.4]	0.8	0.066
NfL (ng/mL)	245.4 [123.8, 591.4]	362.0 [180.7, 718.5]	0.072	0.143
Death in Hospital	5 (6.8%)	16 (33%)	<0.001	0.689
GOSE 6 months	4.0 [3.0, 7.0]	3.0 [1.0, 4.0]	0.004	0.654

[1] Mean (SD); n (%); Median [Q1, Q3]

[2] Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

[3] Standardized Mean Difference

Table 2. Demographic, clinical, and biomarker information by Radiographic neuroworsening groups. Comparison between patients in the unchanging, changing but not worsening, and changing and worsening groups.

Characteristic	Unchanging N = 93[1]	Changing but not worsening N = 10[1]	Changing and worsening N = 16[1]	p-value[2]	Effect size[3]
Age at admission	45.0 (21)	43.4 (20)	48.3 (22)	0.8	<0.001
Gender				0.7	0.08
Male	68 (73%)	6 (60%)	12 (75%)		
Female	25 (27%)	4 (40%)	4 (25%)		
Race				>0.9	0.14
Asian	4 (4.3%)	0 (0%)	1 (6.3%)		
Black	8 (8.6%)	1 (10%)	2 (13%)		
Native American or Native Alaskan	2 (2.2%)	0 (0%)	0 (0%)		
Pacific Islander or Native Hawaiian	1 (1.1%)	0 (0%)	0 (0%)		
White	71 (76%)	8 (80%)	13 (81%)		
Other	4 (4.3%)	0 (0%)	0 (0%)		
Decline to answer	3 (3.2%)	1 (10%)	0 (0%)		
Ethnicity				0.14	0.167
Hispanic or Latinx	7 (7.5%)	2 (20%)	0 (0%)		
Non-hispanic, non-Latinx	84 (90%)	7 (70%)	16 (100%)		
Decline to answer	2 (2.2%)	1 (10%)	0 (0%)		
Admission GCS Score	7 [6, 10]	7 [3, 9]	8 [6, 9]	0.7	<0.001
GCS Motor	7 [7, 8]	7 [6, 7]	7 [7, 7]	<0.001	0.002
Injury Severity Score	22 [16, 29]	26 [26, 33]	26 [18, 37]	0.13	<0.001
Craniotomy				0.2	0.295
No	11 (41%)	0 (0%)	4 (40%)		
Yes	16 (59%)	6 (100%)	6 (60%)		
Intracranial Monitoring				0.011	0.282
No	60 (65%)	3 (30%)	5 (31%)		

Characteristic	Unchanging N = 93[1]	Changing but not worsening N = 10[1]	Changing and worsening N = 16[1]	p-value[2]	Effect size[3]
Yes	33 (35%)	7 (70%)	11 (69%)		
Head Abbreviated Injury Score (HAIS)	4.0 [3.0, 4.0]	5.0 [4.0, 5.0]	4.0 [3.5, 5.0]	0.009	0.001
Rotterdam	3.0 [3.0, 3.0]	4.0 [4.0, 5.0]	4.0 [3.0, 5.0]	<0.001	0.013
GFAP (pg/mL)	2,128.2 [643.3, 5,733.9]	11,715.0 [6,581.0, 44,200.0]	5,359.6 [2,823.9, 17,250.0]	<0.001	0.002
Ang-1 (pg/mL)	9,569.8 [5,631.5, 14,473.0]	10,952.2 [6,520.0, 12,322.1]	7,294.9 [5,510.1, 10,931.4]	0.7	<0.001
Ang-2 (pg/mL)	2,715.8 [2,035.2, 3,516.6]	2,284.7 [1,642.4, 2,534.2]	2,442.7 [1,744.9, 3,345.7]	0.2	<0.001
Syndecan-1 (pg/mL)	2,787.8 [2,289.1, 3,992.3]	2,415.9 [2,172.5, 4,392.4]	3,339.8 [2,079.0, 4,061.8]	0.8	<0.001
Serum Amyloid A (ng/mL)	2,212.6 [1,231.0, 5,297.5]	4,156.5 [1,331.9, 9,475.9]	2,896.7 [1,518.7, 4,522.0]	0.7	<0.001
ICAM-1 (pg/mL)	411,767.4 [321,757.9, 573,364.6]	403,940.7 [324,144.2, 448,700.1]	403,878.7 [322,935.8, 525,450.6]	0.5	<0.001
VCAM-1 (pg/mL)	894,911.5 [693,576.0, 1,097,281.3]	748,339.2 [586,193.9, 1,074,604.5]	817,800.4 [676,301.9, 1,021,095.7]	0.3	<0.001
vWF (ng/mL)	3,662.8 [2,844.0, 5,150.4]	2,629.0 [2,209.5, 4,395.0]	3,270.3 [2,785.1, 5,712.3]	0.4	<0.001
APOA (ng/mL)	1,967,785.3 [1,518,913.1, 2,499,161.4]	1,760,932.3 [1,074,527.9, 2,153,239.3]	1,388,558.7 [1,147,028.7, 2,010,323.8]	0.057	<0.001
IL-6 (pg/mL)	21.9 [12.6, 35.1]	26.2 [20.2, 55.9]	24.2 [14.1, 110.2]	0.3	<0.001
TNF-a (pg/mL)	8.9 [5.9, 15.3]	9.2 [4.7, 13.7]	6.4 [5.0, 8.3]	0.11	<0.001
CRP (ng/mL)	1,727.5 [850.0, 4,062.4]	2,082.6 [1,744.9, 5,786.1]	3,135.7 [942.6, 4,712.8]	0.6	<0.001
NfL (ng/mL)	233.4 [134.9, 586.9]	655.7 [589.2, 1,652.8]	484.7 [232.9, 787.0]	0.008	0.001
Death in Hospital	11 (12%)	5 (50%)	4 (25%)	0.006	0.294
GOSE 6 months	4.0 [3.0, 7.0]	1.0 [1.0, 4.0]	3.0 [1.0, 5.0]	0.083	<0.001

Characteristic	Unchanging N = 93[1]	Changing but not worsening N = 10[1]	Changing and worsening N = 16[1]	p-value[2]	Effect size[3]
[1] Mean (SD); n (%); Median [Q1, Q3]					
[2] Kruskal-Wallis rank sum test; Fisher's exact test					
[3] Partial eta-squared or epsilon squared for continuous; Cramer's V for categorical					

Figures

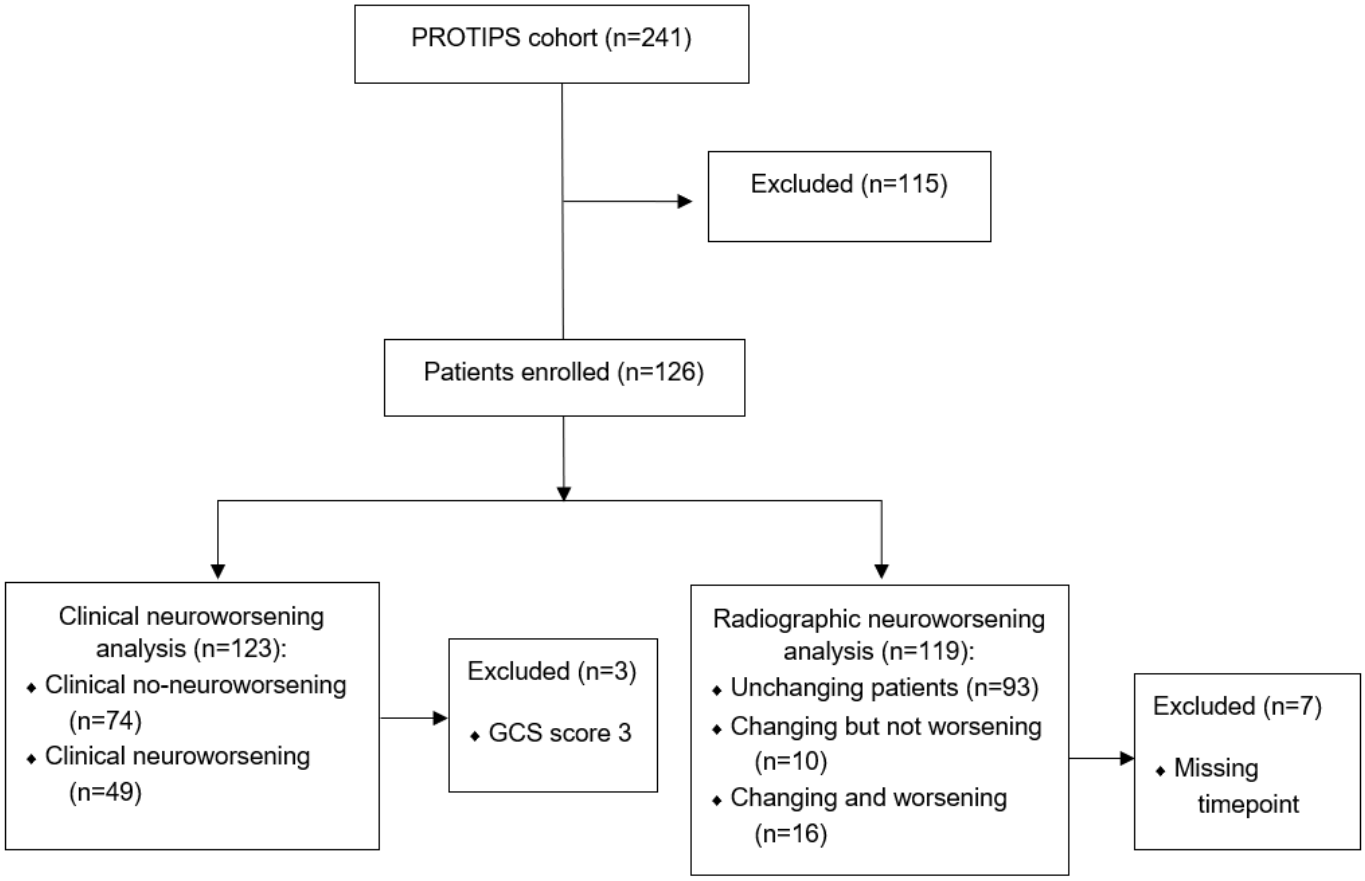

Figure 1

CONSORT Flow diagram of included subjects.

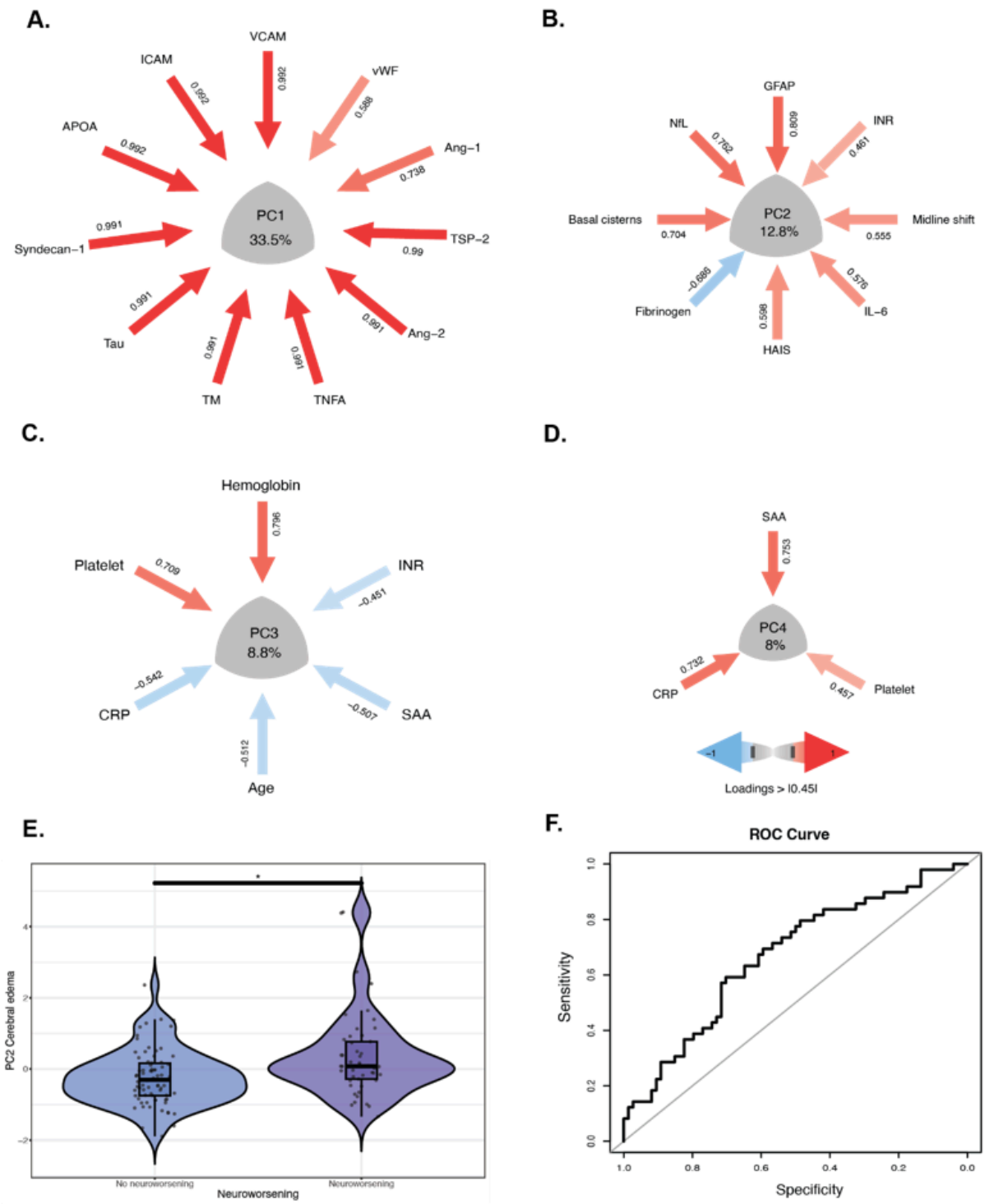

Figure 2

Principal component analysis on GCS neuroworsening. PCA identified 4 endophenotypes that explained 63% of the total variance accounted for in the PROTIPS cohort. **A)** PC1 was labeled as neurovascular-endothelial. **B)** PC2 represented a cerebral edema endophenotype. **C)** PC3 was labeled as pauci-inflammatory. **D)** PC4 was assigned pro-inflammatory state endophenotype. **E)** Cerebral edema endophenotype (PC2) is associated with neuroworsening (p = 0.002, d = 0.60). **F)** The supervised logistic

regression model on PC2 demonstrated modest discriminatory ability (AUC= 0.71) (71%), indicating limited but non-random separation between neuroworsening and no-neuroworsening groups.

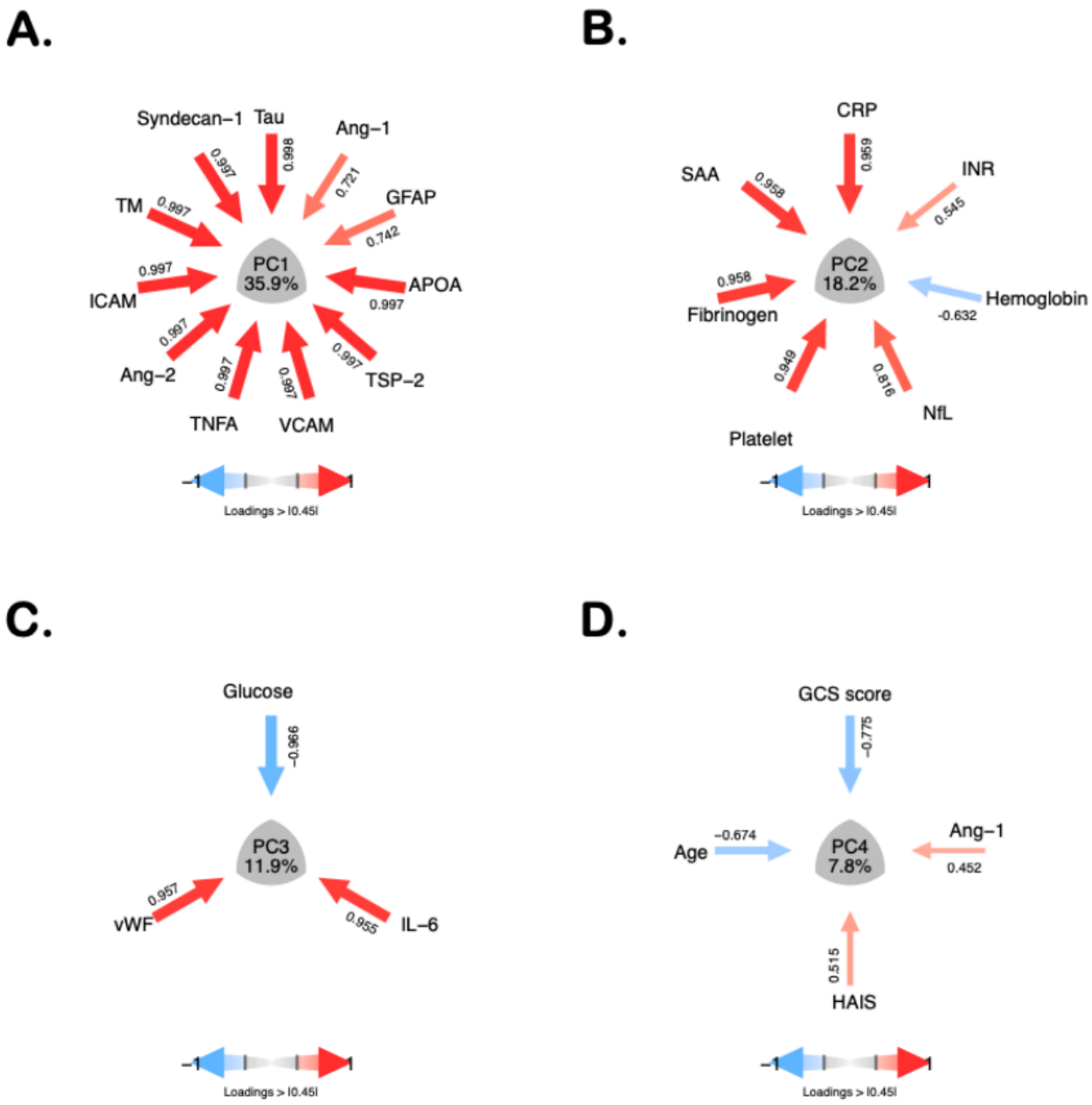

Figure 3

Radiographic neuroworsening PCA identified 4 endophenotypes using the PROTIPS dataset. **A)** PCA1 explained 35.9% of the variation accounted for and was assigned a “vascular injury and edema” endophenotype **B)** PC2 accounted for 18.2% of the total variance and was classified as “pro-inflammatory” endophenotype. **C)** PC3 explained 11.9% of the variation and was classified as “metabolic/endothelial activation” endophenotype. **D)** PC4 accounted for 7.8% of the variation and was assigned an “angiogenic” endophenotype.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- STROBEChecklistNW.docx
- GCSNWSupplementaryText.docx
- HinsonS1.docx
- HinsonS2.docx
- HinsonS3.docx