Received: 3 December 2021 | Revised: 3 May 2022 | Accepted: 2 June 2022

DOI: 10.1002/jts.22868

ISTSS International Society for Traumatic Stress Studies WILEY

RESEARCH ARTICLE

Machine learning prediction of posttraumatic stress disorder trajectories following traumatic injury: Identification and validation in two independent samples

Carissa W. Tomas[1,2] | **Jacklynn M. Fitzgerald**[3] | **Carisa Bergner**[2] | **Cecilia J. Hillard**[4] | **Christine L. Larson**[5] | **Terri A. deRoon-Cassini**[2,6]

[1]Division of Epidemiology, Institute for Health and Equity, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

[2]Comprehensive Injury Center, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

[3]Department of Psychology, Marquette University, Milwaukee, Wisconsin, USA

[4]Department of Pharmacology and Toxicology and Neuroscience Research Center, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

[5]Department of Psychology, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin, USA

[6]Department of Surgery, Division of Trauma and Acute Care Surgery, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

Correspondence
Carissa W. Tomas, Institute for Health and Equity, Medical College of Wisconsin, 8701 W Watertown Plank Road, Milwaukee, WI 53226.
Email: ctomas@mcw.edu

Abstract

Due to its heterogeneity, the prediction of posttraumatic stress disorder (PTSD) development after traumtic injury is difficult. Recent machine learning approaches have yielded insight into predicting PTSD symptom trajectories. Using data collected within 1 month of traumatic injury, we applied eXtreme Gradient Boosting (XGB) to classify admitted and discharged patients (hospitalized, $n = 192$; nonhospitalized, $n = 214$), recruited from a Level 1 trauma center, according to PTSD symptom trajectories. Trajectories were identified using latent class mixed models on PCL-5 scores collected at baseline, 1–3 months posttrauma, and 6 months posttrauma. In both samples, nonremitting, remitting, and resilient PTSD symptom trajectories were identified. In the admitted patient sample, a unique delayed trajectory emerged. Machine learning classifiers (i.e., XGB) were developed and tested on the admitted patient sample and externally validated on the discharged sample with biological and clinical self-report baseline variables as predictors. For external validation sets, prediction was fair for nonremitting versus other trajectories, areas under the curve (AUC = .70); good for nonremitting versus resilient trajectories, AUCs = .73–.76; and prediction failed for nonremitting versus remitting trajectories, AUCs = .46–.48. However, poor precision (< .57) across all models suggests limited generalizability of nonremitting symptom trajectory prediction from admitted to discharged patient samples. Consistency in symptom trajectory identification across samples supports prior studies on the stability of PTSD symptom trajectories following trauma exposure; however, continued work and replication with larger samples are warranted to understand overlapping and unique predictive features of PTSD in different traumatic injury populations.

Each year, over 30,000,000 people are treated for traumatic injuries in hospital emergency departments (EDs) in the United States (Cairns et al., 2018). Despite these high injury volumes, relatively little is known about which factors in the acute period following traumatic injury confer risk or resilience for negative long-term psychological outcomes (Martino et al., 2020). Posttraumatic stress disorder (PTSD) is one of the most common long-term psychological outcomes of trauma exposure, with 10%–40% of survivors developing PTSD (Watkins et al., 2018). Although empirically supported interventions that are provided early after trauma exposure significantly reduce the risk of PTSD development (Watkins et al., 2018), there are currently no validated methods of predicting which individuals have the highest risk.

One barrier to PTSD risk prediction is the inherent heterogeneity of the disorder. Heterogeneity exists in both the symptoms that occur in individuals and in the trajectories of symptoms that manifest over time (Galatzer-Levy et al., 2018). Through the use of machine learning techniques, at least three symptom trajectories have been consistently identified across several trauma samples: chronic, also referred to as nonremitting; recovery, also called remitting; and resilient (Galatzer-Levy et al., 2018). The chronic symptom trajectory consists of clinically elevated or significant symptoms that begin shortly after trauma exposure and persist over time. The recovery or remitting trajectory consists of early clinically significant symptoms that return to baseline after a short period. Finally, the resilient trajectory consists of low or no symptom presentation after trauma exposure, with symptoms that never reach a clinical threshold over time. With accurate and reliable factors differentiating nonremitting individuals, targeted therapeutic interventions could be more efficiently delivered to those at greatest risk (Schultebraucks & Galatzer-Levy, 2019).

Prior work utilizing machine learning approaches to identify and predict PTSD symptom trajectories have demonstrated promising results. For example, in a sample of patients who had been discharged from the ED following traumatic injury, Galatzer-Levy et al. (2017) compared the predictive utility of a combination of biological, clinical, and self-reported variables in differentiating nonremitting and remitting symptom trajectories (Galatzer-Levy et al., 2017). The results demonstrated robust prediction of trajectories, especially when biological and clinical variables were used; however, this approach was not validated in an independent sample. More recently, in discharged (Schultebraucks et al., 2020) and admitted traumatic injury patient samples (Schultebraucks et al., 2021), predictive models using data routinely collected in the ED following injury have shown robust performance. In both studies, the discriminative accuracy of nonremitting versus all other trajectories was excellent and was validated in a second sample (Schultebraucks et al., 2020, 2021). Despite the promise of PTSD trajectory prediction, there is no consensus regarding the specific variables that are routinely predictive, as there are differences in the variables included in the analysis and timing of data collection after traumatic events. Still, a closer review of this work suggests some consistency with regard to predictors, including, but not limited to, prior trauma history, current PTSD symptoms, age, gender, cortisol levels, pain, and injury severity (Galatzer-Levy et al., 2014, 2017; Karstoft et al., 2015; Schultebraucks et al., 2020, 2021).

The current study offers a unique demonstration of external validation in this line of PTSD trajectory prediction and extends prior work in important ways. First, to evaluate if accurate prediction would generalize across samples of traumatically injured adults, we built a predictive model using admitted (i.e., hospitalized) patients and externally validated its performance in a sample of patients discharged from the ED (i.e., nonhospitalized). Such an approach is needed to study predictors of heterogenous PTSD trajectories that encompass the whole traumatic injury population. Marrying the findings for admitted versus discharged patients is essential for understanding whether unique patient characteristics that may influence respective risk versus recovery exist in either sample. Furthermore, we included both biological and clinical self-report measures related to PTSD risk in the predictive models. Although biological predictors are intrinsically of interest given the known pathophysiology of PTSD (Sherin & Nemeroff, 2011), self-report measures provide individual-level experiences crucial to understanding the heterogeneity of the disorder (Ozer et al., 2003). Thus, we included serum concentrations of cortisol and the two primary endocannabinoids: *N*-arachidonoylethanolamine (AEA) and 2-arachidonoylglycerol (2-AG). Preclinical studies and evolving human work support the hypothesis that endocannabinoid signaling is recruited by stress exposure and influences risk regarding the development of mood disorders, anxiety disorders, and chronic pain following traumatic injury (deRoon-Cassini et al., 2020; Fitzgerald et al., 2021).

Considering these factors, our goal was to more explicitly understand if PTSD symptom trajectories and their predictive factors would generalize across traumatically injured civilian samples. In the current study, we first identified trajectories of PTSD symptoms using three assessment points up to 6 months after traumatic injury in one admitted sample and one discharged sample recruited from an urban Level 1 trauma center. Based on consistent trajectory identification in the current literature, we expected to identify nonremitting, remitting, and resilient trajectories in both samples. Sample characteristics within symptom trajectories are also described. In addition, to build upon the previous work in this field, we utilized

acute biological and clinical self-report data to build a predictive model (i.e., extreme gradient boosting; XGB) to classify PTSD symptom trajectory membership in the admitted patient sample, then validate classification performance in the discharged patient sample. A machine learning approach was taken, as this method has demonstrated promising predictive utility in prior studies, and the results of the current study could be benchmarked against previous work. Given the variability of predictor sets in the literature, we assumed an exploratory approach to identifying key predictors in the current study.

METHOD

Participants and procedure

Study on Trauma and Resilience (STAR 1.0; Sample A)

The first sample was from a longitudinal study aimed at using biospecimens, genetics, clinical data, and self-report measures to identify posttrauma risk factors for PTSD development in adults (Fitzgerald et al., 2021; Geier et al., 2019; Weis et al., 2022). Enrolled participants were admitted to the hospital after a traumatic injury (70.3% nonassaultive, 29.6% assaultive). Participants completed three study visits: in-hospital (Time 1 [T1]; $M = 2.5$ days, range: 0–10 days) and 1 month (Time 2 [T2]; $M = 42.28$ days, range: 32–73 days) and 6 months postinjury (T3; $M = 192.2$ days, range: 156–255 days). A total of 278 participants completed T1, 78 completed T2, and 174 completed T3; it is important to note that the T2 assessment was added midway through study recruitment.

Imaging Study on Trauma and Resilience (iSTAR; Sample B)

The second sample was derived from another large longitudinal study, independent from Sample A, that was designed to identify acute posttrauma risk factors for PTSD development in adults using biospecimen, self-report measures, cognitive and behavioral assessments, and neuroimaging (i.e., Bird et al., 2021; Webb, Weis, Huggins, Fitzgerald, et al., 2021; Webb, Weis, Huggins, Parisi, et al., 2021; Weis et al., 2021, 2022). Enrolled participants were recruited after being discharged from the ED after a single-incident traumatic injury (86.0% nonassaultive, 14.0% assaultive trauma) and completed seven study visits. However, the current study utilized only the first three time points: 2–3 weeks (T1; $M = 16.3$, range: 3–30 days), 3 months (T2; $M = 95.6$, range: 72–147 days), and 6 months postinjury (T3; $M = 183.7$, range: 147–240 days). In total, 245 participants completed T1, 214 completed T2, and 227 completed T3.

Both samples were from the same general pool of trauma patients treated at the same urban Level 1 trauma center; however, the samples are independent of one another in that they were recruited at nonoverlapping time intervals using different inclusion and exclusion criteria. See Table 1 for sample characteristics and Supplementary Table S1 for full inclusion criteria, exclusion criteria, and recruitment attrition details for both samples. Both studies were approved by the Medical College of Wisconsin Institutional Review Board.

Thirty-three clinical self-report and biological variables shown to be relevant to PTSD prediction were collected at baseline and considered for prediction across both samples (see Supplementary Table S2 for a description of all variables). Predictor values and labels were recoded for consistency between samples. Participants with greater than 30% missing data for baseline variables were excluded from the analyses (Sample A: $n = 0$, Sample B: $n = 4$). Next, two variables with more than 30% missing data, due to data collection added mid-study, were dropped (i.e., smoking status and the Injured Trauma Survivor Screen score). Predictive mean matching of remaining missing data was done after trajectory identification and before trajectory prediction. The T1 variables retained for analysis are described in the Measures section.

Measures

Demographic characteristics

Participants self-reported their age, sex, race, highest education level completed, employment status ("Are you currently employed or in school?" Yes / No), insurance status (insured / not insured), and home address. Participant address was geocoded and used to derive area deprivation index (ADI) rankings from publicly available data maintained by the University of Wisconsin School of Medicine and Public Health (Kind & Buckingham, 2018). Census block group rankings were derived from the 2014–2018 National American Community Survey (Singh, 2003). National ADI rankings are percentile scores representing 17 variables with scores ranging from 0 (*most advantaged*) to 100 (*most disadvantaged*; Singh, 2003).

Biological characteristics

Heart rate (HR), blood pressure (BP), height (m), and weight (kg) were measured. Body mass index (BMI) was

TABLE 1 Sample characteristics

Variable	Sample A ($n = 192$)				Sample B ($n = 214$)				p
	M	*SD*	*n*	%	*M*	*SD*	*n*	%	
Age (years)	42.17				33.75	10.71			< .001
Gender									< .001
Male			136	70.9			96	44.8	
Female			56	29.1			118	55.1	
Race									< .001
American Indian or Alaska Native				< 5.0[a]					
Asian								< 5.0[a]	
Black or African American			82	42.7			134	62.6	
White			92	47.9			51	23.8	
More than one			16	8.3			25	11.6	
Injury type									< .001
Assaultive			57	29.6			32	14.0	
Nonassaultive			135	70.3			182	86.0	
T1 ISS[b]	10.30	6.02			0.89	2.31			< .001
Days since injury at T1[b]	2.55	1.69			17.62	5.94			< .001
PCL-5 score									
T1[b]	18.91	17.33			28.22	18.79			< .001
T2[c]	20.00	18.74			25.67	18.43			.016
T3[d]	20.40	20.68			21.80	19.70			.480

Note: PCL-5, PTSD Checklist for *DSM-5*; ISS, Injury Severity Score; T1, Time1; T2, Time 2; T3, Time 3;
[a]Small subsample sizes for select racial groups reported as < 5.0% to ensure participant anonymity.
[b]Sample A: at hospital following injury, Sample B: 2–3 weeks postinjury.
[c]Sample A: 1 month postinjury, Sample B: 3 months postinjury.
[d]6 months postinjury for both samples.

calculated as kg/m^2. In both samples, venous blood was obtained once at T1. Concentrations of cortisol and the endocannabinoids (i.e., AEA and 2-AG) were determined as described previously (Fitzgerald et al., 2021) and are detailed in the Supplementary Materials. Due to logistical constraints of the study protocols, blood sample timing could not be standardized but was recorded. Regression analysis indicated there was no association in either sample between blood draw time and cortisol or endocannabinoid levels, (see Supplementary Materials and Figure S1); therefore, time of blood draw was not included in further analyses.

Clinical self-report measures

Injury and pain characteristics

Mechanism of injury and Injury Severity Score (ISS; Baker et al., 1974) were obtained from trauma registrar data in the ED. Participants reported loss of consciousness ("yes" or "no") during their injury, difficulty sleeping since injury ("yes" or "no"), sleep duration in the last 24 hr (in hr), and if they had any history of formal psychiatric diagnosis or treatment ("yes" or "no"). Marijuana use was determined from a urine drug screen at T1. The Visual Analogue Scale for Pain (Holdgate et al., 2003) was used to assess physical pain severity. Participants were asked to rate pain using a numbered line with labels ranging from 0 (*no pain*) to 10 (*worst possible pain*).

Dissociation

The Peritraumatic Dissociation Event Questionnaire (PDEQ; Marmar et al., 1997) was administered to assess dissociation symptoms during the injury event. Participants were asked to rate each of 10 items on a 5-point Likert scale, with total scores of 0–50 and higher scores indicating higher levels of dissociative symptoms. The PDEQ has been validated and has demonstrated good psychometric properties (Tichenor et al., 1996). In the current sample, Cronbach's alpha values were .90 for Sample A and .85 for Sample B.

Trauma exposure

Participants completed the Life Events Checklist (LEC; Gray et al., 2004; Weathers, Blake, et al., 2013), which assesses the occurrence of 17 major life events one may have experienced, witnessed, or learned about happening

to someone close to them. There are numerous validated ways to score the LEC (Weis et al., 2022), and the current study used each of the sum totals of experienced, witnessed, or learned about events. Cronbach's alpha for experienced, witnessed, and learned about scores were .59, .73, and .80, respectively, in Sample A and .67, .79, and .86 in Sample B.

Depression, anxiety, and stress symptoms

The Depression, Anxiety, and Stress Scale (DASS; P. F. Lovibond & Lovibond, 1995; S. H. Lovibond & Lovibond, 1995) is a validated 21-item self-report measure consisting of seven items in each of three subscales. Scores are summed within subscales to evaluate symptoms of depression, anxiety, and stress, respectively, with higher scores indicating more symptoms. Cronbach's alpha values for the Depression, Anxiety, and Stress subscales were .89, .81, and .85, respectively, in Sample A and .90, .83, and .89 in Sample B.

PTSD symptoms

To identify PTSD trajectories, symptoms based on diagnostic criteria in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*; American Psychiatric Association, 2013) were assessed using the 20-item PTSD Checklist for *DSM-5* (PCL-5; Weathers, Litz, et al., 2013) across T1, T2, and T3 within each respective sample. Participants were asked to rate each item on a scale of 0 (*not at all*) to 4 (*extremely*), with total scores ranging from 0–80 and higher scores indicating more severe symptom levels. Subscale scores for the four *DSM-5* symptom clusters (i.e., intrusions, avoidance, negative alterations in cognition and mood, and alterations in arousal and reactivity) can also be calculated by summing scores for related items. Proposed clinical thresholds in other traumatic injury samples suggest a total score of 30 indicates probable PTSD (Geier et al., 2019). To ensure appropriate trajectory identification, only participants with PCL-5 data for at least two of the three assessment points were retained (Sample A: $n = 192$, Sample B, $n = 214$). Symptom cluster scores from the PCL-5 are reported alongside total severity scores in Table 2, though only total severity scores were used in the analysis. Cronbach's alpha values were .92 and .94, respectively, for Sample A and Sample B.

Data analysis

Adherence to reporting standards was assessed with a rubric combining published guidelines, satisfying the Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement (see Supplementary Materials for the checklist; Collins et al., 2015). All code for statistical analysis is included in the Supplementary Materials.

PTSD trajectory identification

PTSD symptom (i.e., PCL-5) trajectories were fit separately for Sample A and Sample B using latent class mixed modeling (LCMM) with the *lcmm* package in R (Proust-Lima et al., 2017). Per recommended guidelines (van de Schoot et al., 2017), to determine the best class fit of one to six classes, multiple criteria were assessed, including reductions of log-likelihood, Akaike information criterion (AIC), Bayesian information criterion (BIC), sample-adjusted BIC (SABIC), and entropy, as well as theoretical interpretability and parsimony. BIC and entropy metrics were weighted more heavily when selecting the best solution (Nguena Nguefack et al., 2020; van de Schoot et al., 2017). For each sample, a split-half cross-validation method was used to validate class selection (Galatzer-Levy et al., 2017), wherein each sample was randomly split in half, LCMM was recalculated, and fit metrics were reassessed. Linear and quadratic slopes were compared to determine the best trajectory fit and shape. Predictor variables were statistically compared between trajectories within each sample. Continuous variables were compared using analyses of variance (ANOVAs), and categorical variables were compared using Fisher's exact test. The Holm–Bonferroni method was used to correct for multiple comparisons (adjusted $p < .05$; Holm, 1979).

Prediction of PTSD trajectories with baseline data

Three models were built to predict trajectory membership: Model 1, which compared nonremitting class membership versus all other trajectories; Model 2, comparing nonremitting versus resilient class membership; and Model 3, which examined nonremitting versus remitting class membership. First, training and test dataset splits from Sample A were stratified by trajectory to ensure adequate representation of trajectories between splits. Models were developed with a training set of T1 data from 75% of Sample A. Model performance was internally validated using the remaining 25% of data in Sample A and externally validated using the full Sample B dataset. For each model, we evaluated performance using two sets of predictor variables: the full set of

TABLE 2 Posttraumatic stress disorder (PTSD) symptom trajectory characteristics

	Sample A[a] ($n = 192$)								Sample B[a] ($n = 214$)					
	Nonremitting ($n = 19$, 9.9%)		Delayed ($n = 29$, 15.1%)		Remitting ($n = 29$, 15.1%)		Resilient ($n = 115$ (59.9%)		Nonremitting ($n = 40$, 18.7%)		Remitting ($n = 31$, 14.5%)		Resilient ($n = 142$, 66.4%)	
Demographic characteristics	***M***	**%**	***M***	**%**	***M***	**%**	***M***	**%**	***M***	**%**	***M***	**%**	***M***	**%**
Time since trauma (days)	2.73		2.86		2.62		2.43		15.75_{a}		16.78		18.33_{a}	
Age (years)	36.84		35.33_{a}		41.41		44.97_{a}		30.13_{a}		31.49		35.28_{a}	
Female sex		57.8_{a}		34.4		37.9		20.8_{a}		60.0		65.6		51.4
Minority race		84.3_{ab}		93.2_{cd}		49.3_{bc}		37.4_{ad}		70.0		84.4		76.1
Not employed		52.6		44.8		34.4		27.8		42.5_{a}		34.3		21.1_{a}
Educational attainment less than high school		26.3		31.0		20.6		13.0		17.5		3.1		9.8
No insurance		15.7		17.2		13.7		14.7		20.0		3.1		26.7
ADI	74.05_{a}		77.13_{b}		66.10		56.23_{ab}		68.37		73.62		69.67	
Biological characteristics														
BMI (kg/m^2)	28.84		28.93		27.69		28.33		29.92_{a}		36.03_{ab}		30.82_{b}	
HR (beats/min)	84.63		79.65		80.93		77.23		78.47		79.62		77.66	
BP systolic	121.63		128.37		131.06		132.12		126.32		132.28		128.73	
BP diastolic	75.52		78.79		75.51		78.61		77.15		76.03		76.89	
AEA	2.21		2.65		2.52		2.48		1.78		1.54		1.50	
2-AG	569.40		977.94		811.12		621.45		14.23		19.90		18.64	
Cortisol	10.21		11.95		16.76		13.81		5.70		7.59		6.32	

(Continues)

TABLE 2 (Continued)

Demographic characteristics	Sample A[a] (*n* = 192)								Sample B[a] (*n* = 214)					
	Nonremitting (*n* = 19, 9.9%)		Delayed (*n* = 29, 15.1%)		Remitting (*n* = 29, 15.1%)		Resilient (*n* = 115 (59.9%)		Nonremitting (*n* = 40, 18.7%)		Remitting (*n* = 31, 14.5%)		Resilient (*n* = 142, 66.4%)	
	M	%	*M*	%	*M*	%	*M*	%	*M*	%	*M*	%	*M*	%
Self-report clinical characteristics														
Assaultive MOI		47.3_{a}		58.6_{b}		31.0		19.1_{ab}		15.0		28.1		11.9
ISS	9.10		12.17		10.27		10.04		1.25		0.81		0.82	
LOC		5.2		6.8		6.8		8.6		20.0		15.6		16.9
THC		21.0		41.3		24.1		28.6		50.0		43.7		40.8
Sleep disturbance		57.8		34.4		41.3		32.1		85.0_{a}		71.8		58.4_{a}
Sleep (hr/night)	5.84		5.94		6.39		6.04		5.68		5.46		6.47	
Previous psychiatric history		47.3		27.5		31.0		23.4		35.0_{a}		25.0		15.4_{a}
PDEQ	38.63_{abc}		26.62_{ad}		28.04_{be}		17.66_{cde}		32.83_{a}		29.93_{b}		24.03_{ab}	
Pain scale	6.73_{a}		6.89_{b}		7.00_{c}		5.19_{abc}		5.45		5.56_{a}		4.04_{a}	
LEC-5														
Events experienced	5.89_{a}		5.20		5.06		4.38_{a}		6.07_{a}		5.65		4.78_{a}	
Events witnessed	5.05		5.03		4.00		4.41		5.87_{a}		5.03		4.23_{a}	
Events learned about	6.52		7.48		6.10		6.87		8.55		7.40		6.80	
DASS														
Depression	11.79_{abc}		3.55_{a}		5.06_{bd}		2.42_{cd}		19.35_{a}		16.68_{b}		5.23_{ab}	
Anxiety	11.78_{abc}		5.71_{a}		7.00_{bd}		3.34_{cd}		17.25_{a}		16.75_{b}		6.01_{ab}	
Stress	13.78_{abc}		6.51_{a}		7.20_{bd}		4.51_{cd}		21.35_{a}		21.68_{b}		9.01_{ab}	

(Continues)

TABLE 2 (Continued)

Demographic characteristics	Sample A[a] ($n = 192$)								Sample B[a] ($n = 214$)					
	Nonremitting ($n = 19$, 9.9%)		Delayed ($n = 29$, 15.1%)		Remitting ($n = 29$, 15.1%)		Resilient ($n = 115$ (59.9%)		Nonremitting ($n = 40$, 18.7%)		Remitting ($n = 31$, 14.5%)		Resilient ($n = 142$, 66.4%)	
	M	%	*M*	%	*M*	%	*M*	%	*M*	%	*M*	%	*M*	%
PCL-5[a,b]														
T1 total score[a]	55.89_{abc}		19.10_{ade}		35.1_{bdf}		8.65_{cef}		48.50_{a}		49.53_{b}		17.71_{ab}	
T1 Intrusion[a]	16.5_{abc}		5.76_{ade}		9.14_{bdf}		1.69_{cef}		12.80_{a}		13.78_{b}		5.15_{ab}	
T1 Avoidance[a]	6.11_{abc}		2.07_{ade}		3.41_{bdf}		0.74_{cef}		5.17_{a}		5.90_{b}		2.28_{ab}	
T1 Arousal/Reactivity[a]	16.1_{abc}		6.17_{ade}		10.7_{bdf}		2.95_{cef}		13.80_{a}		14.21_{b}		5.97_{ab}	
T1 NCAM[a]	17.2_{abc}		5.10_{ad}		12.0_{bde}		3.27_{ce}		16.72_{a}		15.62_{b}		4.30_{ab}	
T2 total score[a]	46.79_{abc}		36.69_{ade}		19.39_{bdf}		13.49_{cef}		46.05_{ab}		41.75_{ac}		16.92_{bc}	
T2 Intrusion[a]	16.54_{ab}		13.63_{c}		10.55_{ad}		3.48_{bcd}		11.18_{a}		10.46_{b}		3.85_{ab}	
T2 Avoidance[a]	6.90_{a}		5.81_{b}		4.77_{c}		1.65_{abc}		5.06_{a}		4.61_{b}		2.19_{ab}	
T2 Arousal/Reactivity[a]	18.36_{ab}		13.63_{c}		11.11_{ad}		4.64_{bcd}		14.68_{a}		13.07_{b}		5.96_{ab}	
T2 NCAM[a]	19.54_{ab}		16.22_{c}		11.22_{ad}		4.87_{bcd}		15.46_{a}		13.80_{b}		4.37_{ab}	
T3 total score[a]	49.18_{abc}		51.06_{ade}		18.79_{bdf}		8.60_{cef}		55.42_{ab}		16.46_{a}		13.80_{b}	
T3 Intrusion[a]	13.25_{ab}		14.00_{cd}		4.64_{ace}		1.69_{bde}		14.17_{ab}		4.28_{a}		3.46_{b}	
T3 Avoidance[a]	6.18_{ab}		6.35_{cd}		2.48_{ace}		1.07_{bde}		5.42_{ab}		2.25_{a}		1.84_{b}	
T3 Arousal/Reactivity[a]	14.06_{ab}		13.85_{cd}		6.36_{ace}		2.76_{bde}		16.90_{ab}		5.78_{a}		4.90_{b}	
T3 NCAM[a]	15.54_{ab}		16.39_{cd}		6.40_{ace}		2.20_{bde}		18.92_{ab}		4.15_{a}		3.55_{b}	

Note: Group and post hoc comparisons were conducted within each respective sample using analyses of variance (ANOVAs) and Fisher's exact tests. Subscript letters indicate significant post hoc pairwise comparisons conducted within each sample after correction for multiple comparisons (i.e., Holm–Bonferroni correction); there was a significant difference between groups with corresponding superscripts. All measures collected at Time 1 (T1) unless otherwise specified; T1 assessments occurred at the hospital postinjury for participants in Sample A and 2–3 weeks postinjury for Sample B.

ADI, Area Deprivation Index; BMI, body mass index; HR, heart rate; BP, blood pressure; AEA, circulating *N*-arachidonoylethanolamine; 2-AG, 2-arachidonoylglycerol; MOI, mechanism of injury; ISS, Injury Severity Score; LOC, loss of consciousness; THC, marijuana use; PDEQ, peritraumatic dissociative experiences questionnaire; LEC, Life Events Checklist; DASS, Depression, Anxiety, and Stress Scales; PCL-5, PTSD Checklist for *DSM-5*; NACM, negative alterations in cognitions and mood; T2, Time 2; T3, Time 3.

[a]For Sample A, T2 assessments occurred 1 month postinjury, and T3 assessments occurred 6 months postinjury; for Sample B, T2 assessments occurred 3 months postinjury, and T3 assessments occurred 6 months postinjury.

[b]Not included in XGB analysis.

baseline variables and a subset of features selected through recursive feature elimination (RFE) during model development. RFE is a dimension-reduction technique that identifies the smallest subset of features that most accurately contribute to classification; this method is often used to protect overfitting in model development when there is a large set of candidate features (Statnikov et al., 2011). To determine the best feature subset from the Sample A training dataset, a 5 x 10-fold cross-validation (i.e., 10 folds, repeated 5 times) was implemented within the *rfe* function in R (Kuhn, 2008).

All data preprocessing was handled separately within training and test sets to reduce information leakage between datasets (Kuhn & Johnson, 2013). Missing data were imputed using predictive mean matching with 20 imputations via the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011). Convergence of imputations was ensured by examining the parallel streams of imputed means and standard deviations. For context, five of 31 variables in Sample A had more than 10% of data missing, and none of the 31 variables in Sample B were missing more than 10% (see Supplementary Figures S2 and S3 for aggregations of missing data plotted using *VIM* in R; Kowarik & Templ, 2016). Categorical variables were dummy-coded using "one-hot encoding". Finally, baseline PCL-5 scores were removed as a predictor due to their role in defining PTSD trajectories. See Supplementary Figure S4 for correlation heat maps of all study variables of interest.

We used the *xgbLinear* method in *train* from the *caret* package in R (Kuhn, 2008) to build an XGB predictive model. XGBoost is an ensemble method that uses decision trees and gradient descent optimization to minimize errors during training (Chen & Guestrin, 2016). During model development, all numeric variables were centered and scaled, and variables with near-zero variance were removed. A random search was implemented for hyperparameter selection (Bergstra & Bengio, 2012), and models with the highest accuracy dictated the final hyperparameter selection. Hyperparameters in *xgbLinear* include the number of boosting iterations, L2 regularization (λ), L1 regularization (α), and learning rate (η; Kuhn, 2008). To tune parameters during model development, repeated cross-validation (i.e., 5 x 10-fold) was applied to guard against overfitting.

Model performance was assessed by examining accuracy, precision, recall, F1 score, and area under the curve (AUC) metrics. Finally, to evaluate variable contributions more closely in the full predictor set, Shapley additive explanation (SHAP) values were calculated for each predictor using the *xgboost* package in R (Chen & Guestrin, 2016; Lundberg & Lee, 2017). SHAP values evaluate the rank-order importance of predictors to classification performance while controlling for the influence and order of all other features in the model (Lundberg & Lee, 2017).

RESULTS

PTSD symptom trajectories

Based on the LCMM fit metrics, a four-class solution for Sample A and a three-class solution in Sample B were selected. In both samples, the results of the split-half cross-validation supported the respective class solution selected (Supplementary Table S3). See Figure 1 and Supplementary Tables S4 and S5 for the best LCMM solution plots and fit metrics for each sample; see Supplementary Figures S5 and S6 for all fitted solutions. Class solutions in Sample A represent the following qualitative trajectories: nonremitting, delayed, remitting, and resilient. In Sample B, the class solutions represent nonremitting, remitting, and resilient trajectories. The nonremitting, remitting, and resilient trajectories in both samples were named according to their consistencies with trajectories identified in previous work. In Sample A, the delayed trajectory was named due to the pattern of subthreshold symptoms at T1 that exceeded the threshold by T2 and persisted to T3. For replicability, we report descriptive comparisons of the nonremitting versus resilient trajectories (all group comparisons are shown in Table 2). In general, participants in the nonremitting trajectory tended to be in a more unstable socioeconomic position (i.e., unemployed and uninsured) and to report more clinically significant psychological symptoms.

In Sample A, when compared to individuals in the resilient trajectory, those in the nonremitting trajectory tended to be primarily female; identify as a racial and/or ethnic minority; and report a lower socioeconomic position (i.e., higher Area Deprivation Index [ADI]; Kind & Buckingham, 2018), higher levels of peritraumatic dissociation, higher levels of pain, more severe PTSD symptoms at all three time points, and higher levels of depression, anxiety, and stress at T1.

In Sample B, the trajectory comparisons were similar. When compared to participants in the resilient trajectory, those in the nonremitting trajectory were younger and more likely to be unemployed; have previously received a psychiatric diagnosis; have experienced and/or witnessed more lifetime traumatic events; have experienced higher levels of peritraumatic dissociation; have had problems sleeping; report higher PTSD symptoms at all three time

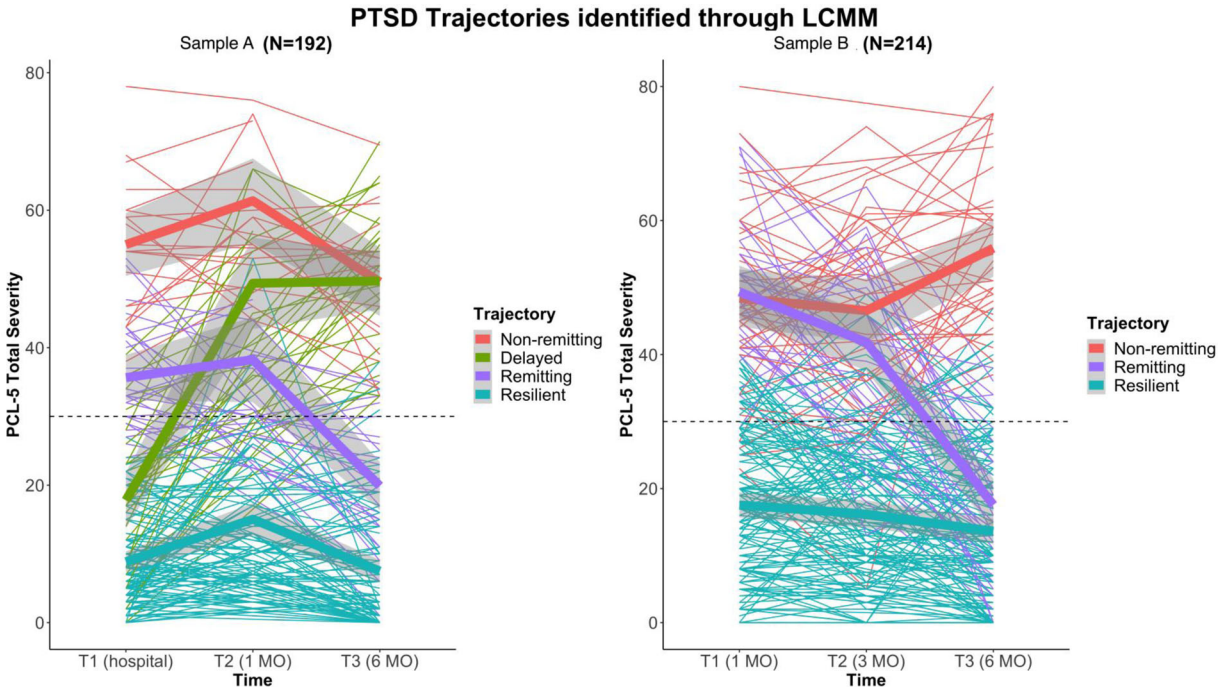

FIGURE 1 Posttraumatic stress disorder (PTSD) trajectories identified using latent class mixed models.
Note: Dashed lines represent PTSD Checklist for *DSM-5* (PCL-5) total severity score proposed clinical cutoff (i.e., 30 or higher) for clinically significant PTSD symptoms. Sample A: nonremitting, $n = 20$; delayed, $n = 29$; remitting, $n = 28$; resilient, $n = 115$. Sample B: nonremitting, $n = 39$; remitting, $n = 34$; resilient, $n = 141$.

points; and report higher levels of depression, anxiety, and stress at T1.

Prediction of symptom trajectories

Model 1: Nonremitting versus all other trajectories

The RFE results indicated there was no smaller subset of baseline variables that best contributed to classifying nonremitting versus all other trajectories; thus, only the results of the model performance when all variables were considered are reported, number of rounds = 59, $\lambda = .04$, $\alpha = .44$, $\eta = 1.44$. When all variables were considered for prediction, performance was good for internal, validation accuracy = .91, precision = .80, recall = .57, F1 = .66, AUC = .82, and fair for external validation, accuracy = .75, precision = .39, recall = .62, F1 = .48, AUC = .70 (Table 3). Performance interpretations were aided by Safari et al. (2016). The top 10 predictors in order of highest importance were depression, peritraumatic dissociation, serum 2-AG, stress, serum AEA, prior indirect trauma exposure (i.e., traumatic events the participant learned about), serum cortisol, race, age, and diastolic blood pressure (Figure 2).

Model 2: Nonremitting versus resilient

Results of the RFE indicated four baseline variables—depression, dissociation, stress, systolic blood pressure—were the smallest set of features contributing to nonremitting versus resilient trajectory membership, number of rounds = 50, $\lambda = .0001$, $\alpha = .0001$, $\eta = 0.3$. When using RFE variables, performance was stronger than for Model 1, with excellent internal validation, accuracy = .93, precision = .85, recall = .85, F1 = .85, AUC = .90, and fair external validation, accuracy = .74, precision = .45, recall = .70, F1 = .54, AUC = .73 (Table 3). When all variables were considered for prediction, model performance was, again, improved compared to Model 1, number of rounds = 29, $\lambda = .00008$, $\alpha = .0001$, $\eta = 1.96$. Performance was good for internal validation, accuracy = .90, precision = .83, recall = .71, F1 = .76, AUC = .83, and fair for external validation, accuracy = .81, precision = .57, recall = .67, F1 = .62, AUC = .76 (Table 3). The top 10 predictors in order of highest variable importance were depression, stress, peritraumatic dissociation, systolic blood pressure, education, prior witnessed traumatic events, age, prior indirect trauma exposure (i.e., traumatic events the participant learned about), race, and serum 2-AG (Figure 2).

TABLE 3 Extreme gradient boosting model performance

Variable and validation	Accuracy	Precision	Recall	F1 score	AUC	AUC 95% CI[a]
Model 1: Nonremitting vs. all other trajectories						
All variables						
Internal	.91	.80	.57	.66	.77	[.57, .92]
External	.75	.39	.62	.48	.70	[.62, .78]
Model 2: Nonremitting vs. resilient						
RFE variables						
Internal	.93	.85	.85	.85	.90	[.74, 1.0]
External	.74	.45	.70	.54	.73	[.65, .80]
All variables						
Internal	.90	.83	.71	.76	.83	[.64, 1.0]
External	.81	.57	.67	.62	.76	[.68, .84]
Model 3: Nonremitting vs. remitting						
RFE variables						
Internal	.64	.66	.57	.61	.64	[.35, .85]
External	.51	.53	.90	.67	.46	[.41, .51]
All variables						
Internal	.71	.80	.57	.66	.71	[.50, .92]
External	.52	.54	.90	.67	.48	[.41, .54]

Note: RFE, recursive feature elimination; AUC, area under the curve.
[a]Calculated using 2,000 stratified bootstrap replicates.

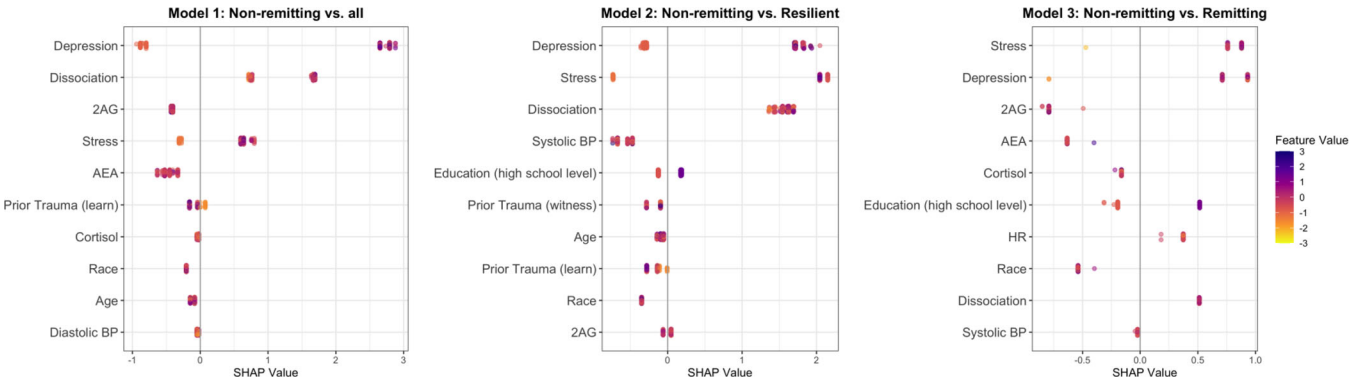

FIGURE 2 Shapley additive explanations (SHAP) values for the top 10 predictors in the full set of predictors for each eXtreme Gradient Boosting model.
Note. Variables are ordered according to classification importance within each respective model=, as calculated from the external validation dataset (i.e., Sample B). SHAP values represent the log-odds probability of predicting membership in the nonremitting PTSD trajectory. Plots provide additional context for model interpretation, although poor overall performance across all models should be noted when interpreting predictor performance. AEA, circulating *N*-arachidonoylethanolamine; 2-AG, 2-arachidonoylglycerol; BP, blood pressure; HR, heart rate.

Model 3: Nonremitting versus remitting

Results of the RFE indicated eight baseline variables—stress, depression, race, cortisol, serum 2-AG, education, serum AEA, and peritraumatic dissociation—were the smallest set of features contributing to trajectory classification, number of rounds = 50, $\lambda = .1$, $\alpha = .1$, $\eta = 0.3$. When using RFE variables, performance was not nearly as strong as Model 1 or 2, with poor internal validation, accuracy = .64, precision = .66, recall = .57, F1 = .61, AUC = .64, and at chance performance (i.e., failed) for external validation, accuracy = .51, precision = .53, recall = .90, F1

= .67, AUC = .4 (Table 3). When all variables were considered for prediction, number of rounds = 80, λ = .0004, α = .01, η = 2.04, performance was fair for internal validation, accuracy = .71, precision = .80, recall = .57, F1 = .66, AUC = .71, and failed for external validation, accuracy = .52, precision = .54, recall = .90, F1 = .67, AUC = .48 (Table 3). The top 10 predictors, in order of importance, were stress, depression, serum 2-AG, AEA, and cortisol, education, heart rate, race, peritraumatic dissociation, and systolic blood pressure (Figure 2).

Overall prediction performance

It is important to note that although Models 1 and 2 demonstrated good performance, particularly in the external validation with respect to accuracy, recall, and AUC, the performance for precision was below chance. Precision in the context of these models is an important metric, as it measures the proportion of positive cases classified as true positives (i.e., nonremitting individuals identified and classified as nonremitting). Given the smaller sample sizes of the nonremitting trajectory (Supplementary Table S6) and that this trajectory is one of the primary clinical groups of interest, this metric should be weighted more heavily to evaluate model performance (Powers, 2011; Saito & Rehmsmeier, 2015). Poor precision (i.e., < .57 in external validation) across all three models suggests overall poor generalization from admitted to discharged patient samples in distinguishing nonremitting individuals from other target groups. For additional context, SHAP value plots (Figure 2) are provided, although poor overall model performance should be noted when interpreting predictor importance.

DISCUSSION

Using samples of admitted and discharged traumatic injury survivors, we assessed the trajectory of PTSD symptoms following injury until approximately 6 months posttrauma. Three subgroups emerged across both samples, characterized as nonremitting, remitting, and resilient. In the admitted patient sample with more severe injuries, we found evidence of a unique fourth subgroup characterized by delayed symptom onset. In addition to mapping trajectories, we predicted subgroup membership using a machine learning approach (i.e., XGB) to test the utility of clinical self-report and biological variables collected shortly after injury (i.e., at T1). Based on internal and external validation, the prediction of nonremitting versus all other trajectories was fair, nonremitting versus resilient trajectories was good, and nonremitting versus remitting trajectories failed.

Although the overall statistical approach of the current study was exploratory, the results provide several important implications for the current understanding of PTSD symptom manifestation following acute injury. With the use of admitted and discharged patient samples, we were particularly well situated to discuss the development of common versus unique trajectories across samples. First, we found consistency in the presence of nonremitting, remitting, and resilient trajectories (Galatzer-Levy et al., 2018), and the proportions of our sample assigned to each trajectory align closely with other trauma center samples (deRoon-Cassini et al., 2010; Galatzer-Levy et al., 2017; Schultebraucks et al., 2020). This consistency suggests that there may be uniformity within the traumatic injury population, whether admitted or discharged, in the relative rates and severity of PTSD symptoms.

Despite such resemblances, we also found that a delayed-onset group was evident in the admitted sample. Supplemental analyses indicated the unique delayed trajectory was erased after controlling for ISS when identifying symptom trajectories (Supplementary Figure S7), suggesting this trajectory is unique to a severely injured subset of traumatic injury patients. Furthermore, individuals in the delayed trajectory (58.6%) experienced more assaultive trauma than those in the remitting (31.0%) or resilient (19.1%) groups. In general, the admitted sample included a more severely injured population, with more reported assaultive trauma; moreover, all participants in this sample required hospitalization, whereas those in the discharged sample did not. When patients are hospitalized or recovering from assault, it is possible that the focus on physical stabilization and recovery is paramount and that only when patients return home are psychological symptoms more evident due to the full realization of how their injuries will affect their well-being, quality of life, and relationships with others. Additionally, it is possible that patients are triggered by reminders and begin to avoid people, places, and situations that remind them of their traumatic experience only after returning home. The presence of this unique subgroup in combination with the generally poor performance of the machine learning classifiers in the discharged sample suggests limited generalizability of symptom trajectory prediction across injury populations. Future work with larger samples is warranted to refine the understanding of overlapping and unique features of admitted and discharged injury populations.

In using machine learning to predict trajectory assignment, several patterns of predictor contributions emerged. It is important to note that standalone predictors cannot be interpreted, as all predictors contribute to model fit

simultaneously. In all models, peritraumatic dissociation, depression, and stress were among the top predictors that contributed to membership in nonremitting versus other trajectories. This finding replicates the well-established association between dissociation and PTSD (Ozer et al., 2003) and underscores the importance of understanding reactions and interactions of symptoms in the peritraumatic period more broadly. The present results showed racial and ethnic minority identity was also a consistent top predictor. Importantly, this finding does not suggest a vulnerability of racial and ethnic minority individuals to poor posttraumatic outcomes (i.e. PTSD) but rather underscores the disparities in trauma exposure and injury prevalence among racial and ethnic groups (Alegría et al., 2013). We also found that circulating concentrations of endocannabinoids (i.e., AEA and 2-AG) at the time of injury were predictive of PTSD risk, which confirms prior work demonstrating high correlations between endocannabinoid levels and trauma outcomes (deRoon-Cassini et al., 2022; Fitzgerald et al., 2021). Although the inclusion of stress-related biomarkers in this study (i.e., cortisol and endocannabinoids) is novel, and the findings support the role that these biomarkers have in prediction, the complex nature of these variables makes the findings preliminary. Future work should examine how biological stress response systems are integrated after trauma exposure, as these stress-related biomarkers may point to possible mechanisms for targeted interventions.

Consistent with previous work, performance was good in classifying nonremitting versus resilient individuals (Schultebraucks et al., 2020). Identifying nonremitting individuals early provides a clear target for preventative therapeutic interventions. Furthermore, the superior performance of this model compared to Models 1 and 3 suggests the stark contrast in the profiles of nonremitting and resilient individuals early after trauma exposure. Distinguishing nonremitting from other at-risk subgroups still presents a challenge, as the classification of nonremitting versus remitting groups in Model 3 failed. Although this comparison is of great clinical interest, these results are not surprising given the time needed for symptoms to diverge. This result highlights the importance of the timing of variable collection for predictive modeling of PTSD symptoms.

This study is not without limitations. First, despite the ability to validate the presence and prediction of PTSD trajectories in two independent traumatic injury samples, there were some distinct methodological differences in data collection between these samples. Although differences in study designs make it difficult to replicate results, they afford the opportunity to unveil important differences regarding the influence of when PTSD symptoms are assessed for mapping PTSD trajectories. Due to constraints of the *lcmm* package in R, we were unable to account for class membership uncertainty. We were also unable to validate latent class selection through likelihood ratio testing and instead relied on BIC, SABIC, and entropy metrics to guide model selection. However, uncertainty modeling and alternative model validation methods are possible within other statistical programs and should be considered for future work and replication. The current study did not account for time of blood draw; however, it is well-known cortisol, as well as endocannabinoids, follow established circadian rhythms and are highly susceptible to acute stress and trauma (Kesner & Lovinger, 2020; Sin et al., 2017). Though time of day has been similarly excluded in previous work (Schultebraucks et al., 2020, 2021), the results regarding cortisol and endocannabinoids in the current study should be interpreted with caution. In addition, although the PCL-5 has strong concordance with clinician-derived assessments of PTSD, we relied solely on a self-report measure of PTSD to determine symptom trajectories. Although the study samples used for validation were independent, further replication across geographically diverse admitted and discharged traumatic injury samples is warranted.

Furthermore, the sample sizes of nonremitting, remitting, and delayed trajectory groups were rather small, which presents a challenge in implementing and interpreting machine learning approaches, as the predicted outcomes are unbalanced. Interestingly, in their review, Galatzer-Levy et al. (2018) noted no association between sample size and the prevalence of trajectories across the literature, suggesting trajectory identification across trauma populations is stable. Still, given that the prevalence of PTSD symptom trajectories will always be unbalanced, the recruitment of larger sample sizes is necessary to clarify meaningful patterns of predictive features. In the context of previous work in this field, though our sample size was on the lower end of the spectrum, we balanced this limitation against the rigor of bias reduction steps employed in the current method (i.e., the maintenance of separate training and test datasets to avoid information leakage, feature selection, hyperparameter search, and cross-validation; Vabalas et al., 2019).

In conclusion, this study was unique in its application of machine learning to evaluate the generalizability of predicting PTSD trajectories from admitted to discharged traumatic injury samples. As machine learning applications are readily employed in medicine and health care for clinical decision-making (Jayatilake & Ganegoda, 2021), it is critical to note that using machine learning for PTSD trajectory prediction is still under development, and there is no consensus on model specification or feature inclusion. Though machine learning can be a powerful tool to distill noisy data, given the inherent heterogeneity of

PTSD symptoms and trauma contexts, as well as the timing of posttrauma assessment, continued work in this field with large and diverse samples is necessary for the replication of results and development of more robust predictive models.

OPEN PRACTICES STATEMENT

Neither of the studies reported in this article was formally preregistered. The data have not been made available on a permanent third-party archive though the script used for analysis can be found in the supplemental material and at the corresponding author's GitHub repository: https://github.com/carissawtomas/PTSD_trajectory_prediction. Requests for data or materials can be sent via email to the lead author at ctomas@mcw.edu.

AUTHOR NOTE

Cecilia Hillard is a member of the Scientific Advisory Boards of Phytecs, Inc. and Formulate Biosciences and has an equity interest in Formulate Biosciences; no other authors have financial disclosures or conflicts of interest to declare.

ORCID

Carissa W. Tomas https://orcid.org/0000-0002-9199-8632
Terri A. deRoon-Cassini https://orcid.org/0000-0002-9485-0625

REFERENCES

Alegría, M., Fortuna, L. R., Lin, J. Y., Norris, F. H., Gao, S., Takeuchi, D. T., Jackson, J. S., Shrout, P. E., & Valentine, A. (2013). Prevalence, risk, and correlates of posttraumatic stress disorder across ethnic and racial minority groups in the United States: *Medical Care*, *51*(12), 1114–1123. https://doi.org/10.1097/MLR.0000000000000007

Baker, S., O'Neill, B., Haddon Jr., W., & Long, W. (1974). The Injury Severity Score: A method for describing patients with multiple injuries and evaluating emergency care. *Journal of Trauma: Injury, Infection, and Critical Care*, *14*(3), 187–196. https://doi.org/10.1097/2F00005373-197403000-00001

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Bird, C. M., Webb, E. K., Schramm, A. T., Torres, L., Larson, C., & deRoon-Cassini, T. A. (2021). Racial discrimination is associated with acute posttraumatic stress symptoms and predicts future posttraumatic stress disorder symptom severity in trauma-exposed black adults in the united states. *Journal of Traumatic Stress*, *34*(5), 995–1004. https://doi.org/10.1002/jts.22670

Cairns, C., Kang, K., & Santo, L. (2018). *National Hospital Ambulatory Medical Care Survey: 2018 emergency department summary tables*. https://www.cdc.gov/nchs/data/nhamcs/web_tables/2018-ed-web-tables-508.pdf

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, *162*(1), 55–63. https://doi.org/10.7326/M14-0697

deRoon-Cassini, T. A., Bergner, C. L., Chesney, S. A., Schumann, N. R., Lee, T. S., Brasel, K. J., & Hillard, C. J. (2022). Circulating endocannabinoids and genetic polymorphisms as predictors of posttraumatic stress disorder symptom severity: Heterogeneity in a community-based cohort. *Translational Psychiatry*, *12*(1), 48. https://doi.org/10.1038/s41398-022-01808-1

deRoon-Cassini, T. A., Mancini, A. D., Rusch, M. D., & Bonanno, G. A. (2010). Psychopathology and resilience following traumatic injury: A latent growth mixture model analysis. *Rehabilitation Psychology*, *55*(1), 1–11. https://doi.org/10.1037/a0018601

deRoon-Cassini, T. A., Stollenwerk, T. M., Beatka, M., & Hillard, C. J. (2020). Meet your stress management professionals: The endocannabinoids. *Trends in Molecular Medicine*, *26*(10), 953–968. https://doi.org/10.1016/j.molmed.2020.07.002

Fitzgerald, J. M., Chesney, S. A., Lee, T. S., Brasel, K., Larson, C. L., Hillard, C. J., & deRoon-Cassini, T. A. (2021). Circulating endocannabinoids and prospective risk for depression in trauma-injury survivors. *Neurobiology of Stress*, *14*, 100304. https://doi.org/10.1016/j.ynstr.2021.100304

Galatzer-Levy, I. R., Huang, S. H., & Bonanno, G. A. (2018). Trajectories of resilience and dysfunction following potential trauma: A review and statistical evaluation. *Clinical Psychology Review*, *63*, 41–55. https://doi.org/10.1016/j.cpr.2018.05.008

Galatzer-Levy, I. R., Karstoft, K.-I., Statnikov, A., & Shalev, A. Y. (2014). Quantitative forecasting of PTSD from early trauma responses: A machine learning application. *Journal of Psychiatric Research*, *59*, 68–76. https://doi.org/10.1016/j.jpsychires.2014.08.017

Galatzer-Levy, I. R., Ma, S., Statnikov, A., Yehuda, R., & Shalev, A. Y. (2017). Utilization of machine learning for prediction of posttraumatic stress: A re-examination of cortisol in the prediction and pathways to non-remitting PTSD. *Translational Psychiatry*, *7*(3), e1070–e1070. https://doi.org/10.1038/tp.2017.38

Geier, T. J., Hunt, J. C., Nelson, L. D., Brasel, K. J., & deRoon-Cassini, T. A. (2019). Detecting PTSD in a traumatically injured population: The diagnostic utility of the PTSD Checklist for *DSM-5*. *Depression and Anxiety*, *36*(2), 170–178. https://doi.org/10.1002/da.22873

Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the Life Events Checklist. *Assessment*, *11*(4), 330–341. https://doi.org/10.1177/1073191104269954

Holdgate, A., Asha, S., Craig, J., & Thompson, J. (2003). Comparison of a verbal numeric rating scale with the Visual Analogue Scale for the measurement of acute pain. *Emergency Medicine Australasia*, *15*(5–6), 441–446. https://doi.org/10.1046/j.1442-2026.2003.00499.x

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Jayatilake, S. M. D. A. C., & Ganegoda, G. U. (2021). Involvement of machine learning tools in healthcare decision-making. *Journal of Healthcare Engineering*, *2021*, 1–20. https://doi.org/10.1155/2021/6679512

Karstoft, K.-I., Galatzer-Levy, I. R., Statnikov, A., Li, Z., Shalev, A. Y., & members of Jerusalem Trauma Outreach and Prevention Study (J-TOPS) group. (2015). Bridging a translational gap: Using machine learning to improve the prediction of PTSD. *BMC Psychiatry*, *15*, 30. https://doi.org/10.1186/s12888-015-0399-8

Kesner, A. J., & Lovinger, D. M. (2020). Cannabinoids, endocannabinoids and sleep. *Frontiers in Molecular Neuroscience*, *13*, 125. https://doi.org/10.3389/fnmol.2020.00125

Kind, A. J. H., & Buckingham, W. R. (2018). Making neighborhood-disadvantage metrics accessible—The neighborhood atlas. *New England Journal of Medicine*, *378*(26), 2456–2458. https://doi.org/10.1056/NEJMp1802313

Kowarik, A., & Templ, M. (2016). Imputation with the *R* package *VIM*. *Journal of Statistical Software*, 74(7). https://doi.org/10.18637/jss.v074.i07

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5). https://doi.org/10.18637/jss.v028.i05

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behavior Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-u

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety & Stress Scales* (2nd ed.). Psychology Foundation.

Lundberg, S., & Lee, S. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4765–4774). MIT Press.

Marmar, C. R., Weiss, D. S., & Metzler, T. J. (1997). The peritraumatic dissociative experiences questionnaire. In *Assessing psychological trauma and PTSD*. (pp. 412–428). The Guilford Press.

Martino, C., Russo, E., Santonastaso, D. P., Gamberini, E., Bertoni, S., Padovani, E., Tosatto, L., Ansaloni, L., & Agnoletti, V. (2020). Long-term outcomes in major trauma patients and correlations with the acute phase. *World Journal of Emergency Surgery*, *15*(1), 6. https://doi.org/10.1186/s13017-020-0289-3

Nguena Nguefack, H. L., Pagé, M. G., Katz, J., Choinière, M., Vanasse, A., Dorais, M., Samb, O. M., & Lacasse, A. (2020). Trajectory modelling techniques useful to epidemiological research: A comparative narrative review of approaches. *Clinical Epidemiology*, *12*, 1205–1222. https://doi.org/10.2147/CLEP.S265287

Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin*, *129*(1), 52–73. https://doi.org/10.1037/0033-2909.129.1.52

Powers, D. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Proust-Lima, C., Philipps, V., & Liquet, B. (2017). Estimation of extended mixed models using latent classes and latent processes: The *R* package *lcmm*. *Journal of Statistical Software*, 78(2). https://doi.org/10.18637/jss.v078.i02

Safari, S., Baratloo, A., Elfil, M., & Negida, A. (2016). Evidence-based emergency medicine; part 5 receiver operating curve and area under the curve. *Emergency*, *4*(2), 111–113.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, *10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Schultebraucks, K., & Galatzer-Levy, I. R. (2019). Machine learning for prediction of posttraumatic stress and resilience following trauma: An overview of basic concepts and recent advances. *Journal of Traumatic Stress*, *32*(2), 215–225. https://doi.org/10.1002/jts.22384

Schultebraucks, K., Shalev, A. Y., Michopoulos, V., Grudzen, C. R., Shin, S.-M., Stevens, J. S., Maples-Keller, J. L., Jovanovic, T., Bonanno, G. A., Rothbaum, B. O., Marmar, C. R., Nemeroff, C. B., Ressler, K. J., & Galatzer-Levy, I. R. (2020). A validated predictive algorithm of post-traumatic stress course following emergency department admission after a traumatic stressor. *Nature Medicine*, *26*(7), 1084–1088. https://doi.org/10.1038/s41591-020-0951-z

Schultebraucks, K., Sijbrandij, M., Galatzer-Levy, I., Mouthaan, J., Olff, M., & van Zuiden, M. (2021). Forecasting individual risk for long-term posttraumatic stress disorder in emergency medical settings using biomedical data: A machine learning multicenter cohort study. *Neurobiology of Stress*, *14*, 100297. https://doi.org/10.1016/j.ynstr.2021.100297

Sherin, J. E., & Nemeroff, C. B. (2011). Post-traumatic stress disorder: The neurobiological impact of psychological trauma. *Dialogues in Clinical Neuroscience*, *13*(3), 263–278. https://doi.org/10.31887/DCNS.2011.13.2/jsherin

Sin, N. L., Ong, A. D., Stawski, R. S., & Almeida, D. M. (2017). Daily positive events and diurnal cortisol rhythms: Examination of between-person differences and within-person variation. *Psychoneuroendocrinology*, *83*, 91–100. https://doi.org/10.1016/j.psyneuen.2017.06.001

Singh, G. K. (2003). Area deprivation and widening inequalities in U.S. mortality, 1969–1998. *American Journal of Public Health*, *93*(7), 1137–1143. https://doi.org/10.2105/AJPH.93.7.1137

Statnikov, A., Aliferis, C. F., Hardin, D. P., & Guyon, I. (2011). *A gentle introduction to support vector machines in biomedicine: Volume 1: theory and methods*. World Scientific. https://doi.org/10.1142/7922

Tichenor, V., Marmar, C. R., Weiss, D. S., Metzler, T. J., & Ronfeldt, H. M. (1996). The relationship of peritraumatic dissociation and posttraumatic stress: Findings in female Vietnam theater veterans. *Journal of Consulting and Clinical Psychology*, *64*(5), 1054–1059. https://doi.org/10.1037/0022-006X.64.5.1054

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLOS ONE*, *14*(11), e0224365. https://doi.org/10.1371/journal.pone.0224365

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3). https://doi.org/10.18637/jss.v045.i03

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S., & Vermunt, J. K. (2017). The GRoLTS-Checklist: Guidelines for reporting on latent trajectory studies. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 451–467. https://doi.org/10.1080/10705511.2016.1247646

Watkins, L. E., Sprang, K. R., & Rothbaum, B. O. (2018). Treating PTSD: A review of evidence-based psychotherapy interventions. *Frontiers in Behavioral Neuroscience*, *12*, 258. https://doi.org/10.3389/fnbeh.2018.00258

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013). *The Life Events Checklist for DSM-5 (LEC-5)*. https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM-5 (PCL-5)*. https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

Webb, E. K., Weis, C. N., Huggins, A. A., Fitzgerald, J. M., Bennett, K. P., Bird, C. M., Parisi, E. A., Kallenbach, M., Miskovich, T., Krukowski, J., deRoon-Cassini, T. A., & Larson, C. L. (2021). Neural impact of neighborhood socioeconomic disadvantage in traumatically injured adults. *Neurobiology of Stress*, *15*, 100385. https://doi.org/10.1016/j.ynstr.2021.100385

Webb, E. K., Weis, C. N., Huggins, A. A., Parisi, E. A., Bennett, K. P., Miskovich, T., Krukowski, J., deRoon-Cassini, T. A., & Larson, C. L. (2021). Neighborhood disadvantage is associated with stable deficits in neurocognitive functioning in traumatically-injured adults. *Health & Place*, *67*, 102493. https://doi.org/10.1016/j.healthplace.2020.102493

Weis, C. N., Huggins, A. A., Miskovich, T. A., Fitzgerald, J. M., Bennett, K. P., Krukowski, J. L., Webb, E. K., deRoon-Cassini, T. A., & Larson, C. L. (2021). Acute white matter integrity post-trauma and prospective posttraumatic stress disorder symptoms. *Frontiers in Human Neuroscience*, *15*, 742198. https://doi.org/10.3389/fnhum.2021.742198

Weis, C. N., Webb, E. K., Damiano, S., Larson, C. L., & deRoon-Cassini, T. A. (2022). Scoring the Life Events Checklist: Comparison of three scoring methods. *Psychological Trauma: Theory, Research, Practice, and Policy*, *14*(4), 714–720. https://doi.org/10.1037/tra0001049

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tomas, C. W., Fitzgerald, J. M., Bergner, C., Bergner, C., Hillard, C. J., Larson, C. L., & deRoon-Cassini, T. A. (2022). Machine learning prediction of posttraumatic stress disorder trajectories following traumatic injury: Identification and validation in two independent samples. *Journal of Traumatic Stress*, *35*, 1656–1671. https://doi.org/10.1002/jts.22868