Derivation and validation of emergency department based risk prediction for posttraumatic stress symptoms six months following traumatic stress exposure

Raphael Kim[1,2], **Gehao Pang**[1,2], Tina Lin[1,2], Andrew Tungate, Samuel A. McLean, MD MPH[1,2,3], Sarah D. Linnstaedt, PhD[1,2]

[1]Institute for Trauma Recovery, University of North Carolina, Chapel Hill, NC, USA. [2]Department of Anesthesiology, University of North Carolina, Chapel Hill, NC, USA. [3]Department of Emergency Medicine, University of North Carolina, Chapel Hill, NC, USA.

ABSTRACT

Posttraumatic stress symptoms (PTSS) are common following trauma exposure. Identification of PTSS-vulnerable individuals would aid in preventative treatment decisions. In the current study, we performed analyses to identify significant clinical predictors and determine their accuracy in predicting PTSS outcomes of trauma. White and Black American men/women (n=1,546) presenting to one of sixteen emergency departments (EDs) within 24 hours of motor vehicle collision (MVC) trauma were enrolled into two independent longitudinal studies and assessed six months following MVC for PTSS (≥33, Impact of Events Scale-Revised). Sociodemographic, pain, general health, event, and psychological/cognitive characteristics were collected in the ED and used in prediction modeling. Ensemble learning methods and Monte Carlo cross-validation were used for feature selection and to determine prediction accuracy. External validation was performed on a hold-out sample (30% of total n) that leveraged the multiple ED enrollment sites. Twenty-five percent (n=394) of individuals reported PTSS six months following MVC. Regularized linear regression was the top performing ensemble learning method and the selected variables showed good reliability in predicting PTSS in the external sample (AUC=0.79+/-0.0017). Top predictors included acute pain severity, expectation of recovery, socioeconomic status, Black vs White ethnicity, and feeling like in a daze. These analyses add to the growing literature indicating that risk for future PTSS can be predicted using characteristics measured from individuals reporting to the ED following trauma. Future studies should aim to replicate these findings in additional trauma cohorts and refine the results for clinical use.

METHODS

Data used in the current study was collected as part of two longitudinal cohort studies of trauma survivors. These two studies enrolled individuals in the Emergency Department (ED) in the immediate aftermath of motor vehicle collision (MVC) trauma, and followed study participants over the course of a year. The first of the two studies enrolled only self-reporting White Americans (June 2011 and June 2014) and the second study enrolled only self-reporting Black Americans (between July 2012 and July 2015). Both sister studies shared the common goal of understanding recovery vs development of adverse posttraumatic neuropsychiatric sequelae following trauma exposure. Sociodemographic, MVC-related pain intensity, MVC-related depressive symptoms, and somatic symptom information was taken at the ED, and at 6 week, 6 month, and 1 year following the MVC.

The White and Black American datasets were cleaned and imputed separately and then merged into a final dataset. Cleaning, variable reduction, and imputation steps were adapted from previously published protocols. Briefly, we first removed variables with >10% missingness. This resulted in a total of 966 variables in the White American dataset and 958 variables in the Black American dataset (of these variables, >90% of them contained complete data). We then used missForest, a nonparametric method of imputation, to impute variables with missing values. Using these complete data, we then scaled continuous covariates ([0,1] range), removed variables with zero or low variance (i.e. those variables in which the fraction of unique values over the sample size was 10% and the ratio of the frequency of the most prevalent value to the frequency of the second most prevalent value was 19) and removed one of any pair of variables in high correlation with each other (i.e. $|r| > 0.75$). Finally, variables not present in both cohorts were removed. A total of 160 variables remained. All cleaning steps were performed using RStudio. Instead of using our two datasets as separate discovery and validation datasets, we opted to combine the White and Black American datasets into one large cohort (facilitated by identical study designs across the two studies) and leverage enrollment study sites as external validation subsets of the large cohort by holding out data from specific enrollment study sites. In order to generate our study site splits, we generated all possible combinations of study sites that could fulfill a 70:30 split between training and test data. Within the training data, we constrained the possible combinations of study sites by three metrics: ratio of White Americans to Black Americans was between 0.45 and 0.55, the ratio of women to men was between 0.45 and 0.55 and every training set had to have at least one study site from each major geographical location (defined as Michigan area, Northern east coast, and Southern east coast). These constraints resulted in 605 different combinations of possible training sets (Supplementary data file X). Due to computational costs, we randomly selected 200 of these splits (as indicated in Supplementary data file X) to perform feature selection and internal validation.

For a given split, we built machine learning models to perform binary classification of PTSS. We compared the performance of regularized logistic regression, random forests, linear support vector machine, and superlearner. In our pipeline, for a given model M, we considered the number of top covariates to use in the model, k, as a hyperparameter to cross validate on with model-specific hyperparameters, α. We selected the most performant (k, α) using the one-standard error rule. To determine the top k covariates, we implemented stability selection from Shah and Samsworth38, which utilizes several rounds of monte carlo cross validation in order to robustly estimate the probability of variable selection. In our procedure, if the support of covariate i was non-zero according to Lasso, then i is considered a signal variable. We calculated the mean and standard error for a variety of performance metrics afterwards.

TABLE 1. Baseline Characteristics of study participants from two longitudinal studies of MVC trauma survivors (n=1,546)

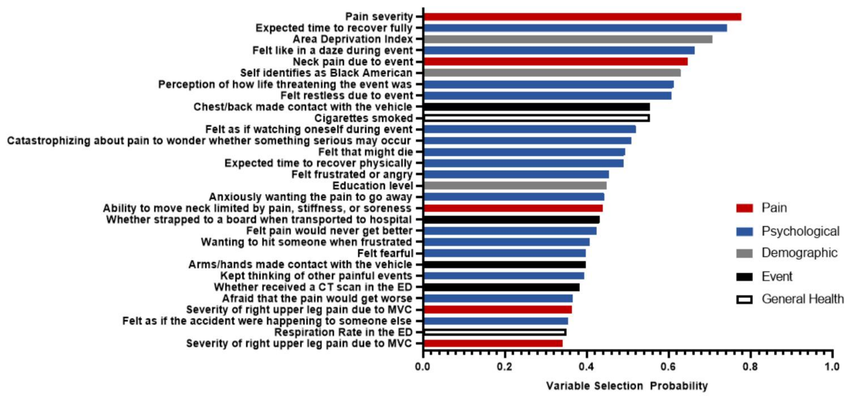
Characteristic		
Ethnicity	**White American**	**Black American**
Participants, n	776	770
Females, n (%)	485 (63%)	495 (64%)
Age, years, mean (SD)	36.3 (13)	35.5 (12.7)
Education, n (%)		
HS or less	169 (22%)	305 (40%)
Some college	304 (39%)	325 (42%)
College	199 (26%)	110 (14%)
Post-college	104 (13%)	30 (4%)
Collision characteristics, n (%)		
Driver	673 (87%)	542 (70%)
Airbag deployed	216 (29%)	228 (30%)
Front end	355 (46%)	339 (44%)
Severe vehicle damage	405 (52%)	412 (54%)
BMI, mean (SD)	27.7 (6.4)	30.0 (7.6)

TABLE 2. Prediction of six-month PTS symptoms using demographic and questionnaire data collected in the ED following MVC. Results presented are the average metrics calculated based on two-hundred stratified splits of the two cohorts into discovery and validation subsets.

	Accuracy	**AUC**	**Sensitivity (True Pos)**	**Specificity (True Neg)**
Regularized Regression				
Training	.78 ± .0013	.86 ± .0013	.78 ± .0030	.78 ± .0030
Internal	.76 ± .0026	.85 ± .0025	.74 ± .0062	.77 ± .0057
External	.70 ± .0027	.79 ± .0017	.71 ± .0054	.70 ± .0050
Random Forest				
Training	.97 ± .0024	.99 ± .00064	.97 ± .0019	.97 ± .0019
Internal	.75 ± .0028	.83 ± .0027	.70 ± .0057	.79 ± .0050
External	.68 ± .0026	.78 ± .0018	.75 ± .0044	.66 ± .0045
SVM*				
Training	.78 ± .0016	.86 ± .0016	.78 ± .0030	.78 ± .0030
Internal	.75 ± .0032	.84 ± .0027	.74 ± .0062	.76 ± .0059
External	.69 ± .0028	.77 ± .0018	.70 ± .0056	.68 ± .0052
Super Learner				
Training	.89 ± .0046	.96 ± .0032	.89 ± .0037	.89 ± .0037
Internal	.75 ± .0030	.84 ± .0030	.74 ± .0050	.76 ± .0048
External	.70 ± .0026	.78 ± .0018	.70 ± .0045	.69 ± .0043

*Linear Support Vector Machines

Figure 1. Shown are the top 30 variables selected via stability selection. Top variable selection probability ranged between 0.78 and 0.34. The colors refer to different categories of variables assessed in the ED.



RESULTS

We averaged the variable importance probabilities from all 200 randomly selected splits to get a holistic sense of the important predictors for PTSD. The top thirty variables are shown in figure 1, where k=30 is chosen by the one standard error rule. The primary variables chosen include pain severity, expected number of days to recover fully, area deprivation index, and self identifies as Black American. Several other psychological, pain, demographic, and event related variables are selected as important predictors. Regarding overall distribution, most of the mass concentrated in these top thirty. External validation and model performance The selected variables are predictive as shown in table 2. We find regularized regression performs the best, with an average AUC of .79 ± .0017 (standard error). Superlearner, random forests, and support vector machines yield external validation metrics closely behind this.

DISCUSSION

We find that our algorithm selects a sparse subset of variables (30) that is predictive of PTSD (.79 AUC) following motor vehicular collision. Important predictors include the expected number of days to recover, pain severity, self identifies as Black American, area deprivation index, and several psychological covariates related to after-effects and perception of event. In the context of the growing literature on PTSD prediction, our work corroborates with the findings of other studies and identifies unique results. Many of these psychological variables were previously shown to be important for PTSD prediction. We also find ethnicity and sociological variables such as ADI, education level, and self identifies as Black American are important. This indicates an ethnic and economic disparity in PTSD outcomes, supporting the need for subgroup focused prediction and treatment5. Interestingly enough, while past work has supported the importance of sex-based differences in PTSS outcomes, we find sex ranks quite low in terms of variable selection probabilities. Furthermore, our clinical prediction tool uniquely identifies ethnicity and pain severity as top predictors for PTSD. While we assessed a comprehensive covariate set and provided rigorous external validation, there are a few key limitations to address. First, the only ethnicities included are Black American and White American. Studies investigating the limitations of Genome-Wide Association Studies have highlighted the importance of performing studies in ethnically diverse datasets. In fact, among what we may characterize as Black American alone should really be distinguished as thirteen different ethnic groups. Second, while the sample size and follow up rate are quite respectable, the sample size still limits the generalizability of our conclusions. Third, there are still covariates we did not directly collect on such as childhood trauma that may be important for PTSD prediction. One could argue that many of the covariates we collected may act as instrument variables for such covariates, but this link is not clear. This work is a promising step towards developing a robust clinical validation tool for PTSD prediction. Our rigorous external validation procedure indicates that we can predict PTSD well using a short and cheap questionnaire (30 questions). Ideally this can be incorporated in the clinic in a manner similar to MDCalc. Regardless at approximately 80% AUC, there is still room for improvement. In future work, we hope to incorporate biological data (genomic and transcriptomic data) and neuroimaging biomarkers in the prediction task. We also hope to combine our results with the results from other studies to improve prediction and reach a consensus on what the important predictors of PTSD are following trauma. Utilizing additional covariates and modalities may assist in obtaining a sparser set of highly predictive features. Future studies should seek to replicate these findings in additional cohorts.

REFERENCES

Dobie DJ, Kivlahan DR, Maynard C, Bush KR, Davis TM, Bradley KA. Posttraumatic stress disorder in female veterans: association with self-reported health problems and

functional impairment. Archives of internal medicine 2004;164:394-400.

Outcalt SD, Kroenke K, Krebs EE, et al. Chronic pain and comorbid mental health conditions: independent associations of posttraumatic stress disorder and depression with pain, disability, and quality of life. Journal of behavioral medicine 2015;38:535-43.

Kessler RC. Posttraumatic stress disorder: the burden to the individual and to society. The Journal of clinical psychiatry 2000.

Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D. Cost of lost productive work time among US workers with depression. Jama 2003;289:3135-44.

Bleich A, Solomon Z. Evaluation of psychiatric disability in PTSD of military origin. Isr J Psychiatry Relat Sci 2004;41:268-76.