

ARTICLE OPEN



Post-traumatic stress and genetic interactions affect tobacco and alcohol use after trauma: findings from a multi-ancestry cohort

Henri M. Garrison-Desany^{1✉}, Cecilia A. Hinojosa², Justin D. Tubbs^{3,4,5}, Jacquelyn L. Meyers⁶, Sarah D. Linnstaedt⁷, Stacey L. House⁸, Francesca L. Beaudoin^{9,10}, Xinming An⁷, Jennifer S. Stevens¹², Thomas C. Neylan¹¹, Gari D. Clifford^{12,13}, Laura T. Germaine^{14,15,16}, Scott L. Rauch^{14,16,17}, John P. Haran¹⁸, Alan B. Storrow¹⁹, Paul I. Musey Jr²⁰, Phyllis L. Hendry²¹, Sophia Sheikh²¹, Christopher W. Jones²², Brittany E. Panches^{23,24}, Jose L. Pascual^{25,26}, Mark J. Seamon^{26,27}, Erica Harris²⁸, Claire Pearson²⁹, David A. Peak^{30,31}, Roland C. Merchant³², Robert M. Domeier³³, Brian J. O'Neil³⁴, Paulina Sergot³⁵, Leon D. Sanchez^{31,32}, Steven E. Bruce³⁶, Steven E. Harte^{37,38}, Samuel A. McLean^{39,40}, Kerry J. Ressler^{16,41}, Karestan C. Koenen⁴² and Christy A. Denckla¹

© The Author(s) 2025

Tobacco smoking and drinking alcohol are common substance use behaviors influenced by both genetic risk and environmental exposures. Traumatic events are highly prevalent, affecting about 70% of people in their lifetime. After trauma, it is unclear what role post-traumatic stress disorder (PTSD) symptoms play in substance use behaviors when accounting for this genetic risk. We used data from the Advancing Understanding of Recovery after trauma (AURORA), which included 2973 participants recruited at emergency departments (EDs) within 72 h of a traumatic event and followed over time. We measured PTSD symptoms via PTSD Checklist for the DSM-5. Tobacco and alcohol consumption as frequency, quantity, and quantity-frequency in the past 30 days. We generated polygenic risk scores with continuous shrinkage for cross-ancestry estimation (PRS-CSx). We tested for main effects between PRS-CSx scores and interactions with PTSD using quasipoisson regression, with week 8 PTSD symptoms and month 6 substance use behaviors after the traumatic event. Tobacco PRS-CSx score increased the risk of tobacco use by 14% (95% CI: 1.01, 1.29, $p = 0.03$), and alcohol PRS-CSx score did not demonstrate consistent associations in the whole cohort (IRR: 1.08, 95% CI: 0.97, 1.19, $p = 0.16$). When stratified by ancestry group, both tobacco (IRR: 1.36, 95% CI: 1.14, 1.61, $p < 0.001$) and alcohol (IRR: 1.24, 95% CI: 1.07, 1.44, $p = 0.005$) PRS-CSx scores were associated with their respective outcomes in the European ancestry subcohort. Participants with lower genetic risk had stronger associations between re-experiencing symptoms and tobacco use, while participants with higher genetic risk demonstrated weaker association between re-experiencing symptoms and tobacco use. A similar pattern was observed for negative alterations in cognition/mood (NACM) symptoms—participants with lower PRS-CSx scores had stronger associations between NACM symptoms with tobacco use, compared to participants with higher PRS-CSx scores. These interactions were both statistically significant, suggesting an antagonistic effect between PRS-CSx scores and PTSD symptoms on tobacco use.

Translational Psychiatry (2025)15:434; <https://doi.org/10.1038/s41398-025-03593-z>

INTRODUCTION

Exposure to traumatic events is common, estimated to affect 70% of people during their lifetime [1], with approximately 10% of trauma-exposed individuals experiencing post-traumatic stress disorder (PTSD) [2]. There are known associations between PTSD and future substance use behaviors [3–5], and 40.1% of people with PTSD are estimated to have a substance use disorder [6]. Substance use, including tobacco smoking and drinking alcohol, is of particular concern because of their associated morbidity and mortality: tobacco and alcohol consumption both increase the risk of heart disease, stroke, and cancers [7–9], and substance use accounts for approximately 14.7% of disability-adjusted life years globally [10]. PTSD is a heterogeneous condition owing both

to biological risks and environmental factors with four major dimensions (avoidance, negative mood/cognition, hyperarousal, and re-experiencing), but to what extent PTSD and its components contribute to substance use risk is unclear. Understanding this interplay, though, could allow for more accurate risk prediction of substance use in trauma-exposed populations, and help tailor effective interventions based on the most salient constructs of PTSD.

Substance use itself is a complex phenomenon, and, similar to PTSD, affected by both biological underpinnings and environmental factors, such as availability of substances, social support, and socioeconomic status. It is estimated from family-based studies that tobacco and alcohol dependency are both about 50%

A full list of author affiliations appears at the end of the paper.

Received: 27 July 2024 Revised: 11 July 2025 Accepted: 1 September 2025

Published online: 24 October 2025

heritable [11–14], though more recent meta-analyses estimated the single nucleotide polymorphism (SNP)-based heritability between 5 to 15% for both tobacco [15] and alcohol use disorders [16]. While the effect of PTSD symptoms on substance use behaviors has been described before as synergistic, and PTSD can interact with other factors, such as sleep problems [17], overall health [18], and childhood adverse experiences [19], to increase the risk of substance use, genetic interactions with PTSD are less documented.

There are three major gaps in prior literature we seek to address. First, most prior studies that have explored gene-environment interactions on substance use behaviors have been candidate gene studies, which do not capture the whole burden of genomic risk [20], and are subject to bias because they only investigated loci with prior documented or hypothesized evidence [21]. Secondly, no study to date has examined interactions with the four components of PTSD, leaving gaps in our understanding whether PTSD globally interacts with the genetic risk of substance use to increase behaviors, or if there are more salient subscales that affect this relationship. Beyond genetics, there is an array of contradictory findings as to which subscale of PTSD symptoms is most associated with substance use behaviors. There has been some evidence that negative cognition and mood symptoms have higher associations with substance use behaviors (regardless of genetic influence) [22, 23], while others have demonstrated associations with re-experiencing symptoms compared to other clusters [24] and hyperarousal [25]. By adjusting for the genetic risk and potential interactions in one of the largest PTSD cohorts to date, we seek to clarify this relationship. And finally, while polygenic risk scores and interactions with environmental exposures have been studied for substance use outcomes, only two to date have examined polygenic risk interactions with trauma [26], but these have been in single ancestry cohorts [27, 28], limiting their generalizability to more diverse populations.

Although historically, substance use genome-wide association studies (GWAS) have been conducted among predominantly European populations [29–31], more recent studies have successfully identified genetic loci associated with substance use in large cohorts representing multiple ancestries [15, 32–34]. Polygenic risk scores (PRS) have yet to fully leverage these multi-ancestry GWASes, and PRS studies for substance use behaviors have primarily been conducted in European-ancestry cohorts [35–37]. Several novel methods have been proposed to leverage shared genetic effects, while properly modeling ancestry-associated differences in linkage disequilibrium and avoiding bias due to population stratification [38–40]. With the development of these statistical methods, we are now better positioned to estimate the genetic burden for substance use in diverse populations, including after trauma. One powerful Bayesian approach, PRS-CSx [41], applies a continuous shrinkage before GWAS summary statistics and LD reference panels from multiple ancestry populations to estimate posterior weights incorporating both shared and ancestry-specific variant effects. PRS-CSx uses a relatively straightforward implementation process to model the joint distribution without the need for a large validation dataset or additional sequencing data, as is necessary for most other multi-ancestry methods [39].

Now, with cross-ancestry statistical methods to grapple with such diverse genetic data, we will leverage prospective data in the first civilian cohort of this scale, with sufficient statistical power to examine separate symptom subscales of PTSD, such as re-experiencing, avoidance, negative cognition/mood, and hyperarousal symptoms. We hypothesize that the polygenic risk for tobacco and alcohol use synergistically interacts with PTSD symptoms to increase substance use behaviors after trauma. We also hypothesize the strength of this association differs by symptom cluster subscale, and negative cognition/mood has the largest association and synergistic interaction with the underlying

genetic risk to increase substance use, due to prior research into the self-medication hypothesis that participants who experience negative mood use the euphoric and dissociative effects of substance use to cope [22].

METHODS

Study sample

The Advancing Understanding of Recovery after trauma (AURORA) study has been previously described and the data is available through the National Institute of Mental Health (NIMH) Data Archive (NDA) [42]. In brief, participants were recruited within 72 h of presenting to an emergency department (ED) after a traumatic event. Qualifying traumatic events included motor vehicle collisions, physical or sexual assault, falls >10 feet, mass casualty incidents, and other events related to experiencing, witnessing, or learning about actual or threatened serious injury, sexual violence, or death, and agreement from a trained research assistant that this is a likely traumatic event. In our sample, only 3.5% ($n = 107$) of participants were enrolled under this “Other qualifying event” category. Participants were followed at one of 29 study sites for one year across 6 time points: ED baseline, week 2, week 8, month 3, month 6, and month 12 post-trauma. At ED baseline, blood biospecimens were collected for testing (as described in the next section). Participants completed self-report study surveys at each follow-up time, including sociodemographic characteristics, trauma-related questionnaires, general mental health questionnaires and substance use behaviors. The target sample size for AURORA was determined via power analyses in the original protocol and 3050 total participants were recruited in the parent study [42].

Biospecimen collection and quality control

Blood biospecimens were collected using the PAXgene tubes at ED sites and frozen at -20°C until shipping to the National Institute of Mental Health (NIMH) Repository and Genomics Resource (NRGR). DNA was then isolated via magnetic bead Chemagic 360 before testing. Biospecimens were processed on Fluidigm SNP trace for quality check. DNA genotyping was conducted using the Illumina Global Screening Array-24 1.0 at the Stanley Center for Psychiatric Research of the Broad Institute. Genotyping quality control included removing rare variants with minor allele frequency (MAF) < 0.005, karyotypic abnormalities determined via B-allele frequency, cryptic relatedness with kinship coefficient > 0.18 (first-degree relatives), removing SNPs with Hardy-Weinberg equilibrium p -value < 1×10^{-8} , removing SNPs with call rates < 98%, and removing participants with sex chromosomes discrepant from reported sex assigned at birth.

Ancestry principal components were estimated using ADMIXTURE [43] within the sample to determine likely ancestry groups. We generated 10 components for global diversity in our sample. The 3 primary ancestry groups were determined via cross-validation within ADMIXTURE, and likely ancestry was assigned based on >90% of the variance explained in the first three respective principal components, corresponding to likely European (EUR), likely African (AFR), and likely admixed American (AMR) ancestry, respectively. A further 10 principal components (PCs) were estimated within these groups. In models with all ancestry groups, 5 PCs were used to allow the model to converge given the small sample of the admixed American population, leading to 5 global PCs, 5 PCs for European ancestry, 5 PCs for African ancestry, and 5 PCs for American Admixed ancestry (20 PCs total). When stratifying European and African ancestry groups, 10 PCs were used.

Polygenic risk scores

We generated polygenic risk scores (PRS) using two methods. First, we conducted PRS for cross-ancestry estimation (PRS-CSx) [38, 41] using all participants. After stratifying by likely ancestry groups, we used PRSice-2 [44] in each group.

Cross-ancestry scores

For PRS-CSx, we used linkage disequilibrium (LD) reference panels from the 1000 Genome Project Phase 3 [45] for European and African ancestry. We used discovery genome-wide association study (GWAS) summary statistics from largest and most current version of the GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN) for tobacco and alcohol [32]. GSCAN included 119,589 AFR ancestry participants, 2,669,029 EUR ancestry participants, and 286,026 AMR ancestry participants. The summary statistics we used for tobacco were calculated in GSCAN as the

average number of cigarettes smoked per week, and for alcohol, the summary statistics from GSCAN were based on the average number of drinks consumed per week. PRS-CSx multi-ancestry meta-analysis scores were used for all subsequent analyses. Global shrinkage parameter ϕ was derived using a fully Bayesian approach given the large population sizes of the discovery GWAS. PRS-CSx was run using Python 3.9 [46].

Ancestry-specific scores

We also stratified by ancestry group and generated ancestry-specific scores using PRSice-2 applied to the corresponding ancestry's summary statistics. A sensitivity analysis was also performed with >50% ancestry individuals, allowing for greater admixture and increasing sample size. Clumping was performed with a window size of 250 kb, an R^2 threshold for clumping of 0.1, and a p -value threshold for clumping of 1.0. The optimal p -value cut-off was determined using thresholding from $5e-08$ to 0.5 that maximized R^2 , with a step-size of 0.00005. PRSice-2 was run using R 4.3 [47]. PRS-CSx and PRSice-2 scores were standardized and converted into deciles to more easily compare between low and high genetic risk groups.

Post-traumatic stress

We conducted further analyses adjusting for post-traumatic stress (PTS) symptoms at baseline, measured by the PTSD Checklist for the DSM-5 (PCL-5) [48], to identify the potential influence of traumatic stress on the relationship between biological risk for substance use and substance use behaviors after trauma. We also calculated the subscale scores by summing items from the relevant symptom cluster subscales for: avoidance, hyperarousal, negative cognition and mood, and re-experiencing symptoms [48]. We report the standardized PCL-5 scores in our models. We report multiplicative interactions and report additive interactions in supplement. For the additive interaction terms, we generated the relative excess risk due to interaction (RERI), the proportion of the overall effect attributable to interaction (AP), and the synergy index (SI) [49].

Substance use outcomes

Substance use was reported for tobacco and alcohol frequency and quantity via self-report. Tobacco consumption included cigarettes, pipes, cigars, chewing tobacco, dip, snuff, hookah, nicotine gum or patches, or e-cigarettes. This information was collected via the PhenX toolkit [50], which asked participants to report the frequency of use in the past 30 days and the average quantity of use in the past 30 days. We then calculated quantity-frequency by multiplying frequency and quantity counts, resulting in the total number of instances of use in the past 30 days. The primary endpoint of interest was 6 months post-trauma, as roughly half of all participants who undergo substance use disorder treatment complete their treatment within 90 days [51, 52], and this was the first endpoint following the traumatic event beyond that time period. Given the skip pattern of the questionnaire and high rates of missingness of healthcare utilization variables available, multiple imputation was not possible to sufficiently estimate and adjust for prior substance use treatment with confidence. Therefore, we used this approach as a means to capture a timepoint where it was unlikely that participants were currently in a treatment program, even if this was not captured in self-report.

Covariates

The primary covariates in these analyses were the ancestry-specific 3 major PC's, an additional top 5 PC's in the combined ancestry models and the sex assigned at birth (male, female, other). While the question was asked, no participants reported "other" sex assigned at birth, therefore responses were binarized as male or female.

Anxiety symptoms were measured using the Patient Reported Outcomes Measurement Information System (PROMIS) Anxiety Bank items [53], and depression symptoms were measured using the PROMIS Depression Short Form [54]. These were included in our initial assessment of covariates and are reported in the Results, however did not demonstrate evidence of confounding results and were not used in the moderation analyses.

In these models, we expected there may be further confounding of the relationship between PTSD symptoms and substance use behaviors based on previous literature [3, 55–58], and adjusted for additional covariates. These included marital status (defined as never married as the reference, married, and grouping separated, divorced, and widowed as "other status" which were dummy coded), race/ethnicity

(defined as non-Hispanic White as the reference, non-Hispanic Black, Hispanic, and non-Hispanic others which were dummy coded) to capture factors of race/ethnicity beyond genetic ancestry, income (defined as \$35,000/year as the reference, \$35,000 to \$75,000, >\$75,000, and "did not report" which were dummy coded), and education (did not complete high school as the reference, completed high school or received GED, completed bachelor's degree, and completed post-graduate degree which were dummy coded), age in years and age² to identify potential generational effects, and finally region of data collection, which was dummy coded (Northeast as the reference, Southeast, Southwest, Midwest, and West coast). All continuous variables (age, depression symptoms, anxiety symptoms, and PTSD symptoms) were standardized by centering to the mean and scaling by the standard deviation.

Missingness

Missingness was assessed descriptively and graphically with the *nanian* package [59]. Missingness was determined to be missing at random by creating a yes/no variable derived from whether PCL-5 score or substance quantity-frequency was missing and we used univariate tests (t-tests for age, depression, anxiety, and PTSD; Chi-squared tests for all categorical variables) to identify whether a given variable was associated with missingness of either the exposure or outcome. No variable was identified as being associated with missing both the exposure and the outcome; therefore, we determined collider stratification bias was unlikely to occur based on our measured variables, and missingness was most likely to be at random. The most notable value that was missing was self-reported income; age, race/ethnicity, gender, and marital status otherwise had high rates of response (>90%). Income was missing among 335 (12.2%) of the sample overall, including participants who selected that they declined to specify their income and participants who skipped the question entirely. We conducted multiple imputation by chained equation to account for missing values, with 20 datasets for 30 iterations each. We pooled results using Rubin's rules [60].

Statistical analyses

Our study sample included 2973 participants with initial genotype data available, and 2747 samples passed quality control. We generated summary statistics of our sample for major covariates and tested for differences of demographic characteristics in substance use with univariate methods. These methods included Pearson correlation coefficients for tobacco and alcohol frequency and age in years, as well as analysis of variance (ANOVA) tests for ancestry, race/ethnicity, income, marital status, and sex assigned at birth categories and substance use frequency outcomes. These were followed with Kruskal-Wallis tests.

We regressed PRS-CSx scores for average cigarette use with respective tobacco frequency and quantity, and average drinks with alcohol frequency and quantity at ED baseline and month 6 post-traumatic event, controlling for 5 ancestry-specific PCs for each of the 3 major ancestry groups, and sex using quasipoisson regression. Quasipoisson regression was used because substance use frequency and quantity were both counts; however, there was evidence of potential overdispersion when fitting an initial Poisson model and testing the dispersion statistic and the quasipoisson distribution relaxes this assumption. We report the regression coefficient and R^2 for the crude and adjusted model. We performed sensitivity analyses independently for each likely ancestry group and repeated this modeling procedure, controlling for 5 global ancestry PCs and 5 ancestry-specific PCs in each of the 3 major ancestry groups (Equation 1). We compared the PRS-CSx ancestry-stratified scores to the PRSice-2 ancestry-stratified scores to determine the best performance.

$$IRR(\text{Substance use quantity} - \text{frequency}) = \beta_{PRS} + \beta_{PTSD} + \beta_{COV}$$

We then adjusted for PTSD symptoms using the PCL-5 scores, controlling for additional sociodemographic covariates that may confound the relationship between traumatic stress and substance use, and conducted interactive analyses between tobacco and alcohol polygenic risk scores, and PTSD symptoms at 8 weeks following the traumatic event. As per recommendations from Keller [61], we further adjusted for all potential interactions between the PRS scores and other covariates included in the model (Equation 2). We did not include PTSD-covariate interaction terms as we experienced model convergence issues when

including that number of interaction terms, and of the models that did converge, none were significant.

$$IRR(\text{Substance use quantity} - \text{frequency}) = \beta_{PRS} + \beta_{PTSD} + \beta_{PRS * PTSD} + \beta_{PRS * COV}$$

PTSD can only be diagnosed after 1 month following a traumatic event per the DSM-5 [62], therefore we selected the earliest timepoint in the AURORA cohort follow-up that was at least 30 days after the index event. We also performed a sensitivity analysis using the highest PCL-5 score in any follow-up period before month 6, when our outcome was ascertained. This was to determine whether PTSD after trauma compounded or demonstrated a threshold effect with pre-existing genetic risk for substance use to influence post-trauma substance use. We then separated these results into the respective subscales for the PCL-5 to identify salient components driving a potential interaction (Equation 3).

$$IRR(\text{Substance use quantity} - \text{frequency}) = \beta_{PRS} + \beta_{PTSD_{subscale}} + \beta_{PRS * PTSD_{subscale}} + \beta_{PRS * COV}$$

Given we tested four subscales across two PRS scores, we used Benjamini-Hochberg correction of the false discovery rate to maintain a family-wise error rate of 0.05, to mitigate inflation of the family-wise error rate while maintaining power in our analyses compared to more conservative approaches such as a Bonferroni correction. We also generated 95% confidence intervals for point estimates.

Sensitivity analysis

As previously mentioned, we generated additive interaction metrics which are reported in supplement. We also generated generalized estimating equations (GEE) models for the repeated measures. We conducted analysis comparing model fit with the autoregressive 1, exchangeable, independent, and unstructured covariance structure. We determined the exchangeable covariance structure had the best fit based on Akaike Information Criterion and Bayesian Information Criterion and convergence. Given, however, that the scores were generated based on summary statistics that used cross-sectional ascertainment of tobacco and alcohol use, and GEEs estimate the slope over the repeated measures, we report this in supplement as the outcome in the GEE model is ultimately different from the summary statistic data the PRS-CSx scores were generated from. And given we did not have a replication sample for our analyses, we generated E-Values to describe the effect size of a potential unmeasured confounder that would be necessary to negate our results.

Ethical approval

This study was determined as non-human subjects research as a secondary data analysis of de-identified pre-existing data by the Harvard Longwood Institutional Review Board. All data from the parent study was collected following written informed consent of participants, and in accordance with the Declaration of Helsinki of 1978 and all relevant updates.

RESULTS

Sample descriptive statistics

In our final analytic sample, 2747 participants had high quality genotyping data available (Table 1). Our sample predominantly self-identified their race/ethnicity as non-Hispanic Black (50.1%, $n = 1376$), and 1461 were identified genetically as belonging to the African ancestry group (55.2%). Further, 35.6% ($n = 977$) self-identified as non-Hispanic White, and 1110 (41.9%) were considered part of the European ancestry group. An additional 78 (2.9%) participants were categorized as likely belong to the American admixed group. Most participants had experienced a motor vehicle collision as the index traumatic event (76.4%, $n = 2247$), and fewer reported physical or sexual assaults (9.8%, $n = 288$). Participants who smoked tobacco compared to those who neither smoked nor drank alcohol less often had completed a bachelor's degree (19.1 vs. 28.1%). There was no statistically significant association between education and alcohol use at baseline ($p = 0.48$); however, there were differences in income, with those drinking alcohol at baseline

having a higher proportion of income $> \$75,000/\text{year}$ ($n = 252$, 14.8%) compared to those who did not consume alcohol or tobacco ($n = 75$, 10.1%) ($p = 0.003$).

PRS performance

The best predictive performance of PRS-CSx scores was consistently observed in the European sample when stratifying by ancestry (Fig. 1). This finding was similar for alcohol scores (Fig. 2), with the best performance in European ancestry yet lower performance in African and American ancestry groups. We present the deciles of PRS-CSx scores in Supplemental Figs. S12 and S13. The tobacco and alcohol scores were significantly associated with their respective outcomes when controlling for ancestry-specific PCs, age, age², sex, and region (Supplemental Tables S1 and S2). For example, tobacco PRS-CSx was associated with greater tobacco quantity-frequency (IRR: 1.17, 95% CI: 1.05, 1.31, $p = 0.005$), and the R² for tobacco quantity-frequency was < 0.01 overall, but 0.04 in the European ancestry subcohort (IRR: 1.36, 95% CI: 1.15, 1.61; $p < 0.001$). These findings were similar for greater quantity-frequency (IRR: 1.13, 95% CI: 1.01, 1.26; $p\text{-value} = 0.03$). The R² for alcohol quantity-frequency was 0.06 for the European ancestry group, but < 0.01 for the total ancestry groups when analyzed together, with a stronger effect in the European ancestry group as well (IRR: 1.24, 95% CI: 1.07, 1.44, $p\text{-value} = 0.005$). We graph the raw scores of the PRS-CSx scores in Supplemental Fig. S11 for reference. PRSice-2 score performance are reported in supplement (Figs. S11 to S8) and demonstrated reduced performance compared to PRS-CSx. Therefore, we conducted primary analyses using PRS-CSx scores.

Associations between PRS and substance use outcomes

We found statistically significant associations with the PRS score and greater tobacco use when adjusting for other mental health conditions and for sociodemographic covariates (Fig. 3). For instance, for tobacco quantity-frequency, the initial PRS effect estimate controlling only for sex, age, and ancestry-specific principal components was 1.17 (95% CI: 1.03, 1.33, $p = 0.02$), while controlling for PTSD, anxiety, depression at baseline, and experiencing a prior traumatic event, the PRS effect estimate was 1.16 (95% CI: 1.02, 1.32, $p = 0.03$). Further adjusting for all additional sociodemographic variables, the estimated PRS effect of 1.14 (95% CI: 1.01, 1.29, $p = 0.03$) for tobacco outcomes did not appreciably deviate from prior estimates. There was not consistent association with alcohol PRS-CSx score and alcohol quantity-frequency outcomes at baseline, and in the final model adjusting for all potential covariates, it was not significant (IRR: 1.08, 95% CI: 0.97, 1.19, $p = 0.16$).

PRS and Post-traumatic stress interaction analyses

In our interaction analyses (Fig. 4), we examined substance use for future associations at month 6 following the traumatic event, based on post-traumatic stress symptoms for week 8 after the traumatic event. These analyses aimed to identify the prospective relationship between pre-existing genetic risk, the addition of PTSD after a traumatic event, and future substance use behaviors. We found evidence of an interactive effect with PTSD for greater tobacco use quantity-frequency (IRR: 0.91, 95% CI: 0.82, 1.00, $p = 0.05$) such that the association between PTSD and tobacco use was weaker among those with high tobacco use PRS. There was no significant main effect of the PRS score (IRR: 1.05, 95% CI: 0.76, 1.44, $p = 0.79$) or joint interactive effect with week 8 PTSD symptoms and the PRS score (IRR: 1.04, 95% CI: 0.94, 1.16, $p = 0.42$) for alcohol quantity-frequency outcomes. We also graphed the predicted tobacco use quantity-frequency at month 6 when stratifying by the 10th and 90th percentile of the tobacco PRS-CSx score in Supplemental Fig. S14.

PRS and Post-traumatic stress subscale interaction analyses

When examining subscale-specific effects, there was a statistically significant interaction between re-experiencing items and the

Table 1. Descriptive statistics and univariate associations of study participants and tobacco and alcohol frequency of use.

	Used neither tobacco nor alcohol at baseline N = 751	Tobacco used at baseline ^a N = 988	P-value	Alcohol used at baseline ^a N = 1705	p-value	Overall N = 2747
Age Mean (Range) ^b	37.26 (18, 74)	34.01 (18, 72)	<0.001	35.75 (18, 74)	0.01	35.98 (18, 74)
Race/ethnicity ^c			0.38		<0.001	
Non-Hispanic Black	417 (55.6%)	539 (54.6%)		812 (47.6%)		1376 (50.1%)
Non-Hispanic White	216 (28.8%)	315 (31.9%)		667 (39.1%)		977 (35.6%)
Hispanic	87 (11.6%)	107 (10.8%)		185 (10.9%)		309 (11.2%)
Non-Hispanic other	29 (3.9%)	25 (2.5%)		36 (2.1%)		74 (2.7%)
Missing	2 (0.3%)	2 (0.2%)		5 (0.3%)		11 (0.4%)
Ancestry group ^{c,d}			0.17		<0.001	
Group 1	255 (35.5%)	366 (38.3%)		742 (45.1%)		1110 (41.9%)
Group 2	24 (3.3%)	20 (2.1%)		49 (3.0%)		78 (2.9%)
Group 3	441 (61.3%)	569 (59.6%)		856 (52.0%)		1461 (55.2%)
Sex assigned at birth ^c			<0.001		<0.001	
Female	512 (68.2%)	518 (52.4%)		1015 (59.5%)		1687 (61.4%)
Male	239 (31.8%)	470 (47.6%)		689 (40.4%)		1059 (38.6%)
Missing				1 (0.1%)		1 (0.04%)
Marital status ^c			<0.001		0.01	
Never married	426 (56.7%)	675 (68.3%)		1069 (62.7%)		1687 (61.4%)
Married	183 (24.4%)	138 (14.0%)		341 (20.0%)		573 (20.9%)
Separated, divorced, widowed	139 (18.5%)	168 (17.0%)		289 (17.0%)		472 (17.2%)
Missing	3 (0.4%)	7 (0.7%)		6 (0.4%)		15 (0.5%)
Education ^c			<0.001		0.48	
Did not complete HS	81 (10.9%)	183 (18.5%)		180 (10.6%)		319 (11.6%)
Completed HS	410 (54.6%)	589 (59.6%)		900 (52.8%)		1484 (54.0%)
Completed bachelor's	211 (28.1%)	189 (19.1%)		485 (28.4%)		745 (27.1%)
Completed graduate	48 (6.4%)	26 (2.6%)		139 (8.2%)		191 (7.0%)
Missing	1 (0.1%)	1 (0.1%)		1 (0.1%)		8 (0.3%)
Income ^c			<0.001		0.003	
<\$35,000/year	439 (58.5%)	661 (66.9%)		899 (52.7%)		1542 (56.1%)
>\$35,000 and <= \$75,000/year	157 (20.9%)	142 (14.4%)		336 (19.7%)		530 (19.3%)
>\$75,000/year	75 (10.1%)	61 (6.2%)		252 (14.8%)		340 (12.4%)
Did not report	79 (10.5%)	124 (12.6%)		218 (12.8%)		335 (12.2%)

^aParticipants who used both tobacco and alcohol at baseline are represented in both columns. There were 708 participants who used both tobacco and alcohol, 280 who used tobacco only, and 997 who used alcohol only.

^bT-test.

^cChi-squared.

^dParticipants' ancestry groups were defined as >90% of the variance explained by the respective ancestry-based principal component. There were 98 participants who were significantly admixed and did not meet this definition.

PRS-CSx score for all types of tobacco use. This interaction was in a similar direction of effect to the larger PCL-5 score with all items. For example, for tobacco quantity-frequency, the interaction effect between re-experiencing items with PRS was 0.90 (95% CI: 0.82, 0.99, $p = 0.03$) when exponentiated to the incidence rate ratio scale, suggesting that the average effect of re-experiencing

symptoms is less among those with high genetic risk, while among those with lower genetic risk for substance use behaviors, re-experiencing symptoms played a larger role in their future substance use. There was an interaction between the negative cognition and mood subscale and the genetic PRS as well, with IRR 0.90 (95% CI: 0.82, 0.99; $p = 0.02$). This suggests a similar direction

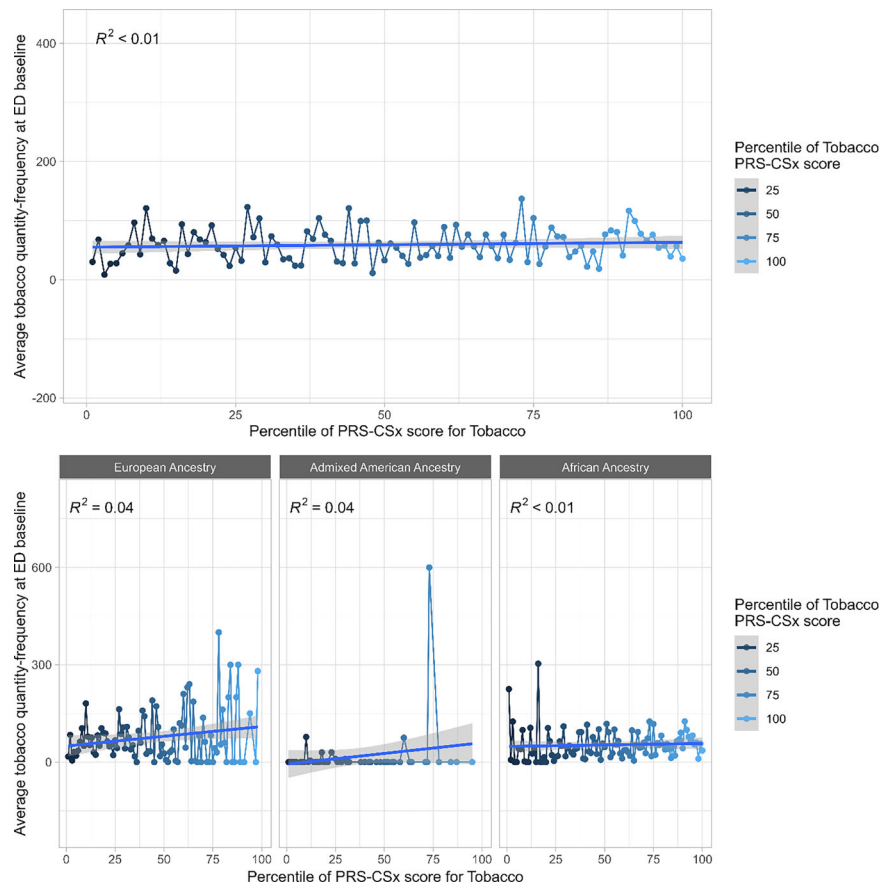


Fig. 1 Percentile of tobacco-based PRS-CSx score and average tobacco use frequency at baseline, for all ancestries and stratified by estimated ancestry group. PRS-CSx scores were first centered and standardized, then converted to deciles for plotting. Ancestry groups were determined using ADMIXTURE software, and thresholds for likely ancestry were set at >0.90 in the factor loadings. The frequency of tobacco use was ascertained in the past 30 days of the emergency department (ED) baseline.

of effect, whereby the average effect of negative cognition and mood symptoms is less among those with high genetic risk (Supplemental Fig. S14). The main effects of PRS and all other subscales were significantly increased tobacco use risk, though no additional interaction was significant (Fig. 4). Alcohol demonstrated a lack of lack of a main effect with the PCL-5 overall and respective subscales in our models (Figs. 3 and 4).

We report the interactions on the additive scale as well, which did not demonstrate significant additive interaction effects for either the RERI, the AP, or the SI (Supplemental Tables S4 and S5).

Sensitivity analyses

We report the generalized estimating equation for tobacco (Supplemental Table S6) and alcohol outcomes (Supplemental Table S7) using the exchangeable correlation structure. There was a significant interaction between the re-experiencing subscale and PRS-CSx (IRR: 0.94, 95% CI: 0.90, 0.99). No interactions with PTSD subscales and the alcohol standardized PRS score were significant. When using the maximum PTSD score in the ED baseline, week 2, week 8, and month 3, re-experiencing (IRR: 0.1, 95% CI: 0.82, 1.00, $p = 0.047$) and negative cognition and mood (IRR: 0.91, 95% CI: 0.83, 1.00, $p = 0.049$) were significant (Supplemental Table S8). There was no significant interaction using the GEE for alcohol outcomes (Supplemental Table S9). With our E-Value analysis (Supplemental Table S10), to negate the significant interaction terms would require an effect size of 1.50 (lower 95% CI limit: 1.21).

We did not have adequate statistical power to identify interaction effects in ancestry group strata, however we did

identify similar main effects. We found an increase in tobacco quantity-frequency among European participants associated with increased PRS-CSx score when controlling for the PCL-5 main and interaction effects (Supplemental Table S11), though there was not a significant relationship between alcohol risk score and alcohol quantity-frequency in the European subsample (Supplemental Table S12). When examining subscales of the PCL-5, there were consistent associations among the European ancestry strata with the tobacco scores increasing tobacco quantity-frequency (Supplemental Table S12), none were statistically significant with the adjusted p -values, however the hyperarousal subscale indicated a suggestive increase (IRR: 1.17, 95% CI: 0.999, 1.36, $p = 0.051$). For the African ancestry strata, the tobacco outcomes are reported in Supplemental Table S13 and were significantly associated with negative cognition/mood symptoms (IRR: 1.20, 95% CI: 1.01, 1.2), and all subscales were associated with alcohol outcomes (for example, hyperarousal had an IRR of 1.28 (95% CI: 1.11, 1.48, $p < 0.001$), but the standardized PRS-CSx score for alcohol use was not significantly associated with quantity-frequency of use (Supplemental Table S14).

DISCUSSION

This study is one of the first studies to investigate the interaction of cross-ancestry polygenic risk scores with PTSD symptoms in relationship to for substance use behaviors after trauma. To our knowledge, this study is also one of the first studies to examine the interaction between specific PTSD subscales and genetic risk on substance use outcomes. Given the high level of co-occurrence

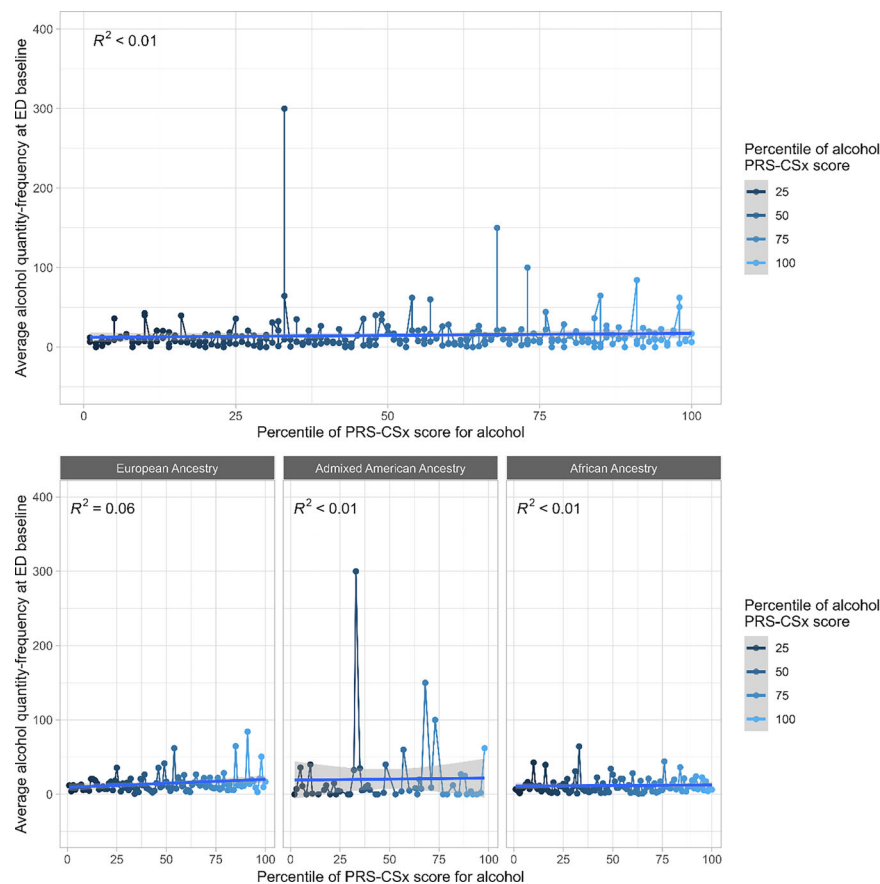


Fig. 2 Percentile of alcohol-based PRS-CSx score and average alcohol use frequency at baseline, for all ancestries and stratified by estimated ancestry group. PRS-CSx scores were first centered and standardized, then converted to deciles for plotting. Ancestry groups were determined using ADMIXTURE software, and thresholds for likely ancestry were set at >0.90 in the factor loadings. The frequency of alcohol use was ascertained in the past 30 days of the emergency department (ED) baseline.

between PTSD and substance use, using longitudinal cohorts to delineate temporal ordering ensures that PTSD symptoms are ascertained before future substance use behaviors. This is paramount for guiding resources and treatment decisions. We demonstrated statistically significant interaction between PTSD symptoms and genetic risk to daily cigarette consumption on post-trauma future daily nicotine product consumption but did not find evidence of an interaction with genetic risk and PTSD symptoms on future alcohol use. Our findings suggest there may be a threshold effect of risk occurring among groups with high genetic risk for tobacco. However, we note that overall prediction of substance use traits in non-European ancestry samples in this investigation was relatively poor. The use of multi-ancestry methods, however, allowed us to examine these associations in a diverse cohort using all participants when controlling for additional principal components and covariates.

While there have been major strides in cross-ancestry methods for PRS, notable challenges remain [39]. Generating and assessing the accuracy of these scores in multiple populations and communities remains a frontier of the field. These challenges are particularly pronounced in complex conditions that have such a notable environmental component, such as substance use, with an estimated heritability between 30 and 70% for nicotine use disorder using twin-based studies [63], and between 50 and 64% for alcohol use disorder using twin-based studies [11]. SNP-based heritability analyses from recent GWAS report lower estimates: a 2024 meta-analysis estimated roughly 5 to 15% heritability of tobacco use disorder phenotypes [15], and while another meta-analysis estimated heritability of alcohol use disorder phenotypes

as 7 to 13% using SNPs [16]. This is further exacerbated by the relatively low predictive ability of current polygenic risk scores, which, even when using more ancestrally homogenous and large cohorts such as the FinnTwin12 study, the PRS explained between 2.5 to 3.5% of the variance for alcohol use disorder [64]. Using multiple ancestry reference groups may boost power for diverse cohorts [40], however we found that performance remained the best in the ancestry group most represented in the discovery GWAS: the European ancestry subcohort. The lack of consistent association in the other groups demonstrates the ongoing limitations of statistical genetics work to leverage multi-ancestry cohorts for PRS estimation. Even using a different method for prediction, PRSice-2, in the ancestry-specific estimates continued to demonstrate low predictive accuracy. This suggests that, despite significant progress in the field, there continue to be challenges based on the limitations of the initial GWAS. This work further highlights the need for greater diversity in GWAS discovery cohorts. In the future, using reference GWAS with large, diverse cohorts, such as All Of Us, could improve PRS prediction.

Our interaction analyses suggest that compounding genetic risk with PTSD symptoms for future tobacco use leads to a lesser increase in substance use than expected from the main effects alone. While further replication of this interaction is necessary, this observation may be due to a thresholding effect in our sample, which was entirely trauma-exposed and reported high rates of overall substance use (e.g., alcohol use was greater than 60%, tobacco use was greater than 30% of the sample). The risk of substance use related to genetic liability may mask the effects of the PTSD symptoms; in other words, individuals are over a

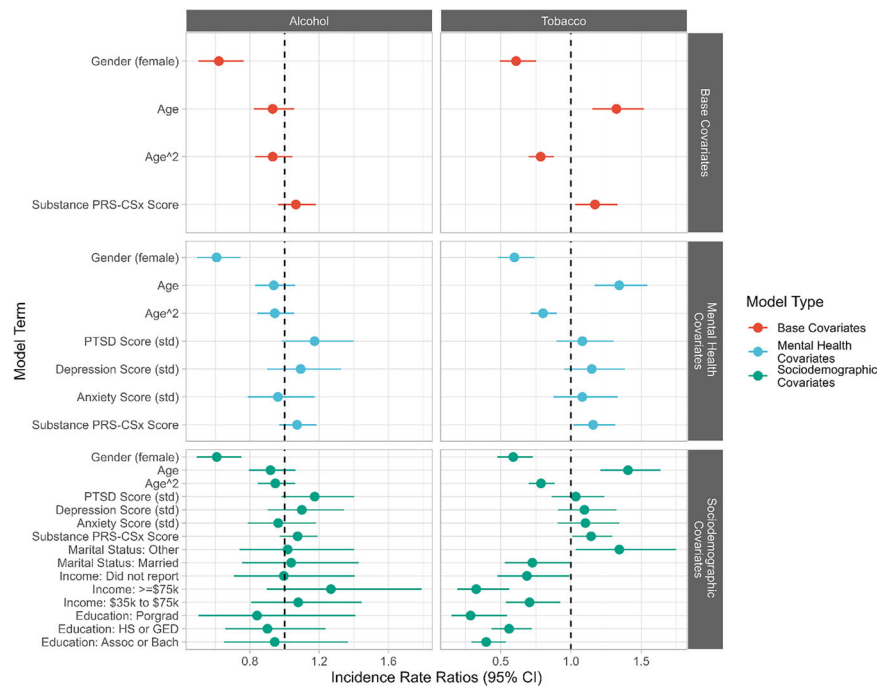


Fig. 3 Associations of polygenic risk scores (PRS) with tobacco frequency and quantity, and alcohol frequency and quantity 6 months after a traumatic event, adjusting for post-traumatic stress and lifetime traumatic events for all ancestries. IRR: Incidence rate ratio as estimated using the quasipoisson model to relax dispersion parameter assumptions, 95% CI: 95% confidence interval. The outcome definition was the number of times one consumed the substance in the past 30 days, multiplied by the quantity consumed on average when one did consume, in the past 30 days. PRS-CSx: Polygenic risk score using continuous shrinkage and cross-ancestry methods; generated with summary statistics from the GSCAN database for average cigarette per week and average alcoholic drinks per week. PTS symptoms: Post-traumatic stress (PTS) was ascertained with the PTSD Checklist for the DSM-5 (PCL-5). Anxiety was ascertained with the PROMIS Anxiety Bank of state and trait anxiety. Depression was ascertained with the PROMIS 8-item short form for Depression. Ancestry-specific principal components were controlled for in all models.

threshold of risk factors, and additional insult of PTSD symptoms does not lead to the same increase as it does for those with low underlying risk. This would be consistent with the diathesis-stress model, whereby individuals with low genetic vulnerability may have increased sensitivity to environmental factors that increase their risk for substance use behaviors [65, 66]. Under such a model, certain participants respond to post-traumatic stress with increases in substance use behaviors, while others do not. In this case, this group vulnerable to PTSD effects is the low genomic risk for substance use. Alternatively, the PRS themselves may capture some PTSD risk, therefore those with high PRS are already at risk of high PTSD after trauma, with less of the measured PTSD explaining changes in tobacco and alcohol consumption. A recent review only found three studies to date directly addressing this effect, with inconclusive evidence and no consistent direction of the interaction effect [26]. Therefore, risk for PTSD symptoms may differ from exposure to prior traumatic events, and future research should investigate how the timing of traumatic events, the number of prior traumatic events, and PTSD symptoms may compound and influence this risk.

The lack of association between PTSD symptoms and alcohol outcomes in our study could be related to the high rate of alcohol consumption at baseline in our sample. Alcohol use disorder has a high prevalence among PTSD-diagnosed patients, ranging from 9.8 to 61.3% [67]. The alcohol PRS was not consistently associated with outcomes when controlling for prior mental health conditions, suggesting that the effect of genetic prediction on alcohol use outcomes may not be independent of these factors. Prior studies, including a systematic review in 2020 [22], have demonstrated there is some evidence of the “self-medication” hypothesis, whereby PTSD symptoms are associated with increased alcohol use. These studies, however, have

methodological limitations, and only a handful of studies reported longitudinal assessments of prior exposure to traumatic events and future alcohol use [4, 68–71]. These studies generally found an increase in alcohol use after trauma, though associations with PTSD symptoms were mixed at times; for example, a study examining alcohol consumption associated with PTSD symptoms after 09/11/2001 found no association with increased use [71]. However, a more recent study with 8 years of follow-up identified that substance use was independent from PTSD symptoms over time, when accounting for other co-morbidities and confounders [72]. Therefore, although we did not demonstrate associations with increased alcohol use, unmeasured confounders which we did not include in our study might have affected the observed relationships. In addition, the various types of traumas experienced by those in the cohort could have been confounders associated with different PTSD symptomatology that we were unable to disentangle given that our cohort had primarily experienced motor vehicle collisions.

When separating by symptom subscales, we found that re-experiencing symptoms and negative cognition/mood subscales demonstrated interactions with the tobacco PRS-CSx score similar to the overall PTSD-PRS interactions. However, hyperarousal and avoidance symptoms did not interact with genetic risk, though the tobacco PRS-CSx score remained consistently associated when accounting for them. These interactions with re-experiencing and negative cognition/mood remained statistically significant when adjusting for co-occurring anxiety and depression symptoms and were negative, compared to positive associations in the main effects. This finding suggests that these two subscales may be independent of other affective symptoms. In prior research, subscale studies conducted using the DSM-4 version of the PTSD Checklist—which included a re-experiencing subscale with

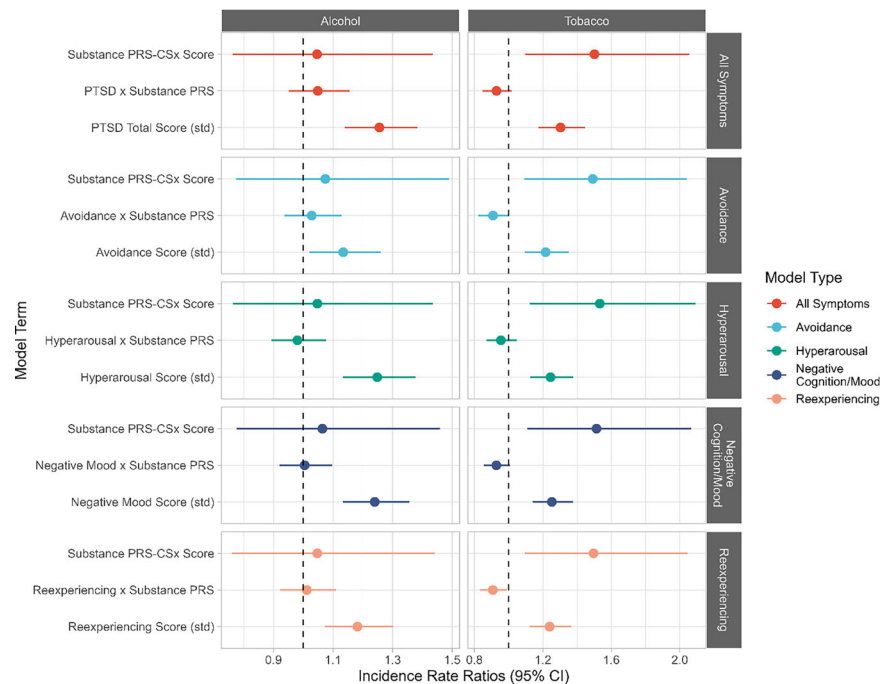


Fig. 4 Main and joint effects of polygenic risk score (PRS) and post-traumatic stress (PTS) symptoms and subscales at 8 weeks post-trauma on tobacco use in the past 30 days at 6 months post-trauma for all ancestries. ¹PRS-CSx: Polygenic risk score using continuous shrinkage and cross-ancestry methods; generated with summary statistics from the GSCAN database for average cigarette per week and average alcoholic drinks per week. ²IRR: Incidence rate ratio as estimated using the quasipoisson model to relax dispersion parameter assumptions, 95% CI: 95% confidence interval. ³PTSD: Post-traumatic stress, symptoms were assessed using the post-traumatic stress disorder (PTSD) symptom checklist for the DMS-5 (PCL-5), with relevant subscales also reported (re-experiencing, negative cognition and mood, hyperarousal, and avoidance). Models were also adjusted with 5 principal components for global ancestry, and 5 principal components calculated within the top 3 ancestry groups, determined by >90% of the genetic variance related to ancestry being explained by that component.

questions worded slightly differently and included the “emotional numbing” subscale—found that emotional numbing was often associated with increased nicotine use and dependence [73, 74], however re-experiencing symptoms were associated with positive and negative reinforcement of smoking behaviors [75]. In one recent study among former smokers exposed to a traumatic event, the investigators found emotional numbing/avoidance was related to cross-sectional assessment of relapse of smoking, but not re-experiencing symptoms [74]. Our findings in a longitudinal cohort may relate to changes in smoking behaviors in the future, rather than prevalence of smoking, and demonstrate how re-experiencing symptoms indicates some other underlying biological vulnerability that interacts with the genetic risk for tobacco consumption.

Our study was limited in several ways: we relied on self-reporting of substance use behaviors. Therefore, we may be subject to misclassification and reporting biases based on recall or social desirability bias [76]. We also used previously described methods for polygenic risk score estimation, which are limited for cross-ancestry populations [39]. While PRS-CSx showed statistically significant and consistent associations with our outcome of interest, suggesting minimal contamination of the score, it underperformed in African ancestry samples compared to European ancestry sample. There also were not enough American Indigenous ancestry samples to accurately estimate risk when stratifying for either PRS-CSx or PRSice-2. Future methods that improve prediction in underrepresented populations may reveal further insights we cannot explore with current methods. We used GSCAN data to build the PRS scores, which specifically quantified cigarette use, however this variable was the closest to the tobacco product frequency and quantity we had available in AURORA. Therefore, our AURORA outcome variable reflected other products

beyond cigarettes, which may bias findings. Although we did find interactions for tobacco outcomes, we may have been underpowered to detect them in alcohol outcomes. Due to the high rates of alcohol use in our sample, there were fewer contrasts available, which may have reduced statistical power. Additionally, our sample had primarily experienced motor vehicle collisions as their index trauma, and other cohorts, such as veteran cohorts or those experiencing physical or sexual assault, may yield different findings. PTSD symptomatology may differ based on the type of trauma experienced, and therefore certain symptom presentations may not be represented in our study. The sample was further skewed by being self-identifying primarily as female and being Black/African-American compared to the general US population. We did not have access to a comparable replication population that had longitudinal data of similar timing available or subscale information with similar measures of tobacco and alcohol quantity-frequency. In the future, replication must be conducted to validate the findings, as in gene-environment research, there are often challenges with replication in additional cohorts. While, to preserve power in our sample, we did not split our data into a training/testing dataset, we did generate E-values for our primary findings to quantify the necessary influence of an unmeasured confounder to negate our results. Finally, there might be residual confounders in our study that we did not consider which could affect our findings.

Our study has several strengths, particularly the longitudinal nature of our data that allowed us to assess future substance use behaviors and trajectories over time. This advantage allowed us to establish temporal ordering of the effect of PTSD symptoms on future substance use behaviors. We also used a cross-ancestry method for polygenic risk scores that showed consistent associations when controlling for co-occurring mental health

conditions and for socioenvironmental factors often related to both substance use scores and substance use behaviors. This observation demonstrates that the PRS largely captured genetic variation rather than an aggregated environmental effect, which may be due to the accounting for ancestry differences. Education was one of our most associated covariates with substance use but it does not diminish the association with the genetic scores in the European-only and all-ancestry cohort. We also report PTS subscale associations, which revealed important differences between different facets of traumatic stress and highlighted particular risks that may affect the risk for substance use. Finally, our sample was a diverse civilian cohort who had all experienced a traumatic event, which likely enables our findings to be more externally valid to general populations.

CONCLUSION

This study demonstrates evidence of genetic and post-traumatic stress interactions on tobacco outcomes and the persistent effect of genetic risk on alcohol outcomes when controlling for other environmental conditions. Our work highlights re-experiencing and negative cognition/mood constructs of PTSD symptoms may be particularly salient in interactions with genetic risk for tobacco consumption and future tobacco outcomes after trauma. These findings show a threshold effect may occur, whereby increased genetic risk and increased PTSD symptoms do not simply aggregate, but rather high-risk individuals may not have further risk conferred after a certain point. Future research may consider different substance use trajectories after traumatic stress, such as differences in remitting or relapsing patterns of use versus incident use, and expand this work to other substances, such as cannabis and opioids.

DATA AVAILABILITY

Data and/or research tools used in the preparation of this manuscript were obtained from the National Institute of Mental Health (NIMH) Data Archive (NDA). NDA is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in mental health. Dataset identifier: Longitudinal Assessment of Post-traumatic Syndromes. AURORA Project #2526. <https://doi.org/10.15154/my7y-qx79>. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDA.

CODE AVAILABILITY

R and Python code used in this analysis upon reasonable request to the corresponding author.

REFERENCES

- Benjet C, Bromet E, Karam EG, Kessler RC, McLaughlin KA, Ruscio AM, et al. The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. *Psychol Med*. 2016;46:327–43.
- Koenen KC, Ratanatharathorn A, Ng L, McLaughlin KA, Bromet EJ, Stein DJ, et al. Posttraumatic stress disorder in the World Mental Health Surveys. *Psychol Med*. 2017;47:2260–74.
- Garrison-Desany HM, Meyers JL, Linnstaedt SD, House SL, Beaudoin FL, An X, et al. Post-traumatic stress and future substance use outcomes: leveraging antecedent factors to stratify risk. *Front Psychiatry*. 2024;15. <https://doi.org/10.3389/fpsy.2024.1249382>.
- Hinojosa CA, Liew A, An X, Stevens JS, Basu A, van Rooij SJH, et al. Associations of alcohol and cannabis use with change in posttraumatic stress disorder and depression symptoms over time in recently trauma-exposed individuals. *Psychol Med*. 2024;54:338–49.
- Brady KT, McCauley JL, Back SE. The comorbidity of Post-traumatic Stress Disorder (PTSD) and substance use disorders. In: el-Guebaly N, Carrà G, Galanter M, Bal-dacchino AM (eds). *Textbook of addiction treatment: international perspectives*. Cham: Springer International Publishing; 2021, pp 1327–39.
- Simpson TL, Rise P, Browne KC, Lehavot K, Kaysen D. Clinical presentations, social functioning, and treatment receipt among individuals with comorbid life-time PTSD and alcohol use disorders versus drug use disorders: findings from NESARC-III. *Addiction*. 2019;114:983–93.
- Falkstedt D, Wolff V, Allebeck P, Hemmingsson T, Danielsson A-K. Cannabis, tobacco, alcohol use, and the risk of early stroke. *Stroke*. 2017;48:265–70.
- Jiang H, Livingston M, Room R, Chenhall R, English DR. Temporal associations of alcohol and tobacco consumption with cancer mortality. *JAMA Network Open*. 2018;1:e180713.
- Gallaway MS. Surveillance for cancers associated with tobacco use — United States, 2010–2014. *MMWR Surveill Summ*. 2018;67. <https://doi.org/10.15585/mmwr.ss6712a1>.
- Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the global burden of Disease study 2010. *PLoS ONE*. 2015;10:e0116820.
- Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*. 2015;45:1061–72.
- Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry*. 2008;65:674–82.
- Agrawal A, Budney AJ, Lynskey MT. The co-occurring use and misuse of cannabis and tobacco: a review. *Addiction*. 2012;107:1221–33.
- Deak JD, Miller AP, Gizer IR. Genetics of alcohol use disorder: a review. *Curr Opin Psychol*. 2019;27:56–61.
- Toikumo S, Jennings MV, Pham BK, Lee H, Mallard TT, Bianchi SB, et al. Multi-ancestry meta-analysis of tobacco use disorder identifies 461 potential risk genes and reveals associations with multiple health outcomes. *Nat Hum Behav*. 2024;8:1177–93.
- Zhou H, Kember RL, Deak JD, Xu H, Toikumo S, Yuan K, et al. Multi-ancestry study of the genetics of problematic alcohol use in over 1 million individuals. *Nat Med*. 2023;29:3184–92.
- Vandrey R, Babson KA, Herrmann ES, Bonn-Miller MO. Interactions between disordered sleep, post-traumatic stress disorder, and substance use disorders. *Int Rev Psychiatry*. 2014;26:237–47.
- Garrison-Desany HM, Childress CP, McConico N, Jarrett BA, Howell S, Glick JL. Substance use patterns among a global sample of transgender and non-binary people during the COVID-19 pandemic. *BMC Glob Public Health*. 2023;1:16.
- Khoury L, Tang YL, Bradley B, Cubells JF, Ressler KJ. Substance use, childhood traumatic experience, and Posttraumatic Stress Disorder in an urban civilian population. *Depress Anxiety*. 2010;27:1077–86.
- Jiang Z, Chen Z, Chen X. Candidate gene-environment interactions in substance abuse: a systematic review. *PLoS ONE*. 2023;18:e0287446.
- Border R, Johnson EC, Evans LM, Smolen A, Berley N, Sullivan PF, et al. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *AJP*. 2019;176:376–87.
- Hawn SE, Cusack SE, Amstadter AB. A systematic review of the self-medication hypothesis in the context of posttraumatic stress disorder and comorbid problematic alcohol use. *J Trauma Stress*. 2020;33:699–708.
- Boland M, Rielage JK, Hoyt T. The power of negative mood in predicting post-traumatic stress disorder and alcohol abuse comorbidity. *Psychol Trauma*. 2018;10:572–5.
- Trautmann S, Schönfeld S, Behrendt S, Schäfer J, Höfler M, Zimmermann P, et al. Associations between lifetime PTSD symptoms and current substance use disorders using a five-factor model of PTSD. *J Anxiety Disord*. 2015;29:93–100.
- Livingston NA, Farmer SL, Mahoney CT, Marx BP, Keane TM. The role of PTSD symptom clusters and criterion in predicting future high-risk drug and alcohol use among returning veteran men and women. *Psychol Serv*. 2022;19:386–95.
- Pasman JA, Verweij KJH, Vink JM. Systematic review of polygenic gene-environment interaction in tobacco, alcohol, and cannabis use. *Behav Genet*. 2019;49:349–65.
- Polimanti R, Kaufman J, Zhao H, Kranzler HR, Ursano RJ, Kessler RC, et al. Trauma exposure interacts with the genetic risk of bipolar disorder in alcohol misuse of US soldiers. *Acta Psychiatr Scand*. 2018;137:148–56.
- Meyers JL, Cerdá M, Galea S, Keyes KM, Aiello AE, Uddin M, et al. Interaction between polygenic risk for cigarette use and environmental exposures in the Detroit Neighborhood Health study. *Transl Psychiatry*. 2013;3:e290.
- Sanchez-Roige S, Fontanillas P, Jennings MV, Bianchi SB, Huang Y, Hatoum AS, et al. Genome-wide association study of problematic opioid prescription use in 132,113 23andMe research participants of European ancestry. *Mol Psychiatry*. 2021;26:6209–17.
- Furberg H, Kim Y, Dackor J, Boerwinkle E, Franceschini N, Ardisino D, et al. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet*. 2010;42:441–7.
- Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, et al. Sequence variants at CHRN3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet*. 2010;42:448–53.
- Saunders GRB, Wang X, Chen F, Jang S-K, Liu M, Wang C, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature*. 2022;612:720–4.

33. Hatoum AS, Colbert SMC, Johnson EC, Huggett SB, Deak JD, Pathak GA, et al. Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders. *Nat Mental Health*. 2023;1:210–23.
34. Xu K, Li B, McGinnis KA, Vickers-Smith R, Dao C, Sun N, et al. Genome-wide association study of smoking trajectory and meta-analysis of smoking status in 842,000 individuals. *Nat Commun*. 2020;11:5302.
35. Chang L-H, Couvy-Duchesne B, Liu M, Medland SE, Verhulst B, Benotsch EG, et al. Association between polygenic risk for tobacco or alcohol consumption and liability to licit and illicit substance use in young Australian adults. *Drug Alcohol Depend*. 2019;197:271–9.
36. Belsky DW, Moffitt TE, Baker TB, Biddle AK, Evans JP, Harrington H, et al. Polygenic risk and the developmental progression to heavy, persistent smoking and nicotine dependence: evidence from a 4-Decade Longitudinal Study. *JAMA Psychiatry*. 2013;70:534–42.
37. Taylor M, Simpkin AJ, Haycock PC, Dudbridge F, Zuccolo L. Exploration of a polygenic risk score for alcohol consumption: a longitudinal analysis from the ALSPAC cohort. *PLoS ONE*. 2016;11:e0167360.
38. Ruan Y, Lin Y-F, Feng Y-CA, Chen C-Y, Lam M, Guo Z, et al. Improving polygenic prediction in ancestrally diverse populations. *Nat Genet*. 2022;54:573–80.
39. Kachuri L, Chatterjee N, Hirbo J, Schaid DJ, Martin I, Kullo IJ, et al. Principles and methods for transferring polygenic risk scores across global populations. *Nat Rev Genet*. 2024;25:8–25.
40. Zhang H, Zhan J, Jin J, Zhang J, Lu W, Zhao R, et al. A new method for multi-ancestry polygenic prediction improves performance across diverse populations. *Nat Genet*. 2023;55:1757–68.
41. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019;10:1–10.
42. McLean SA, Ressler K, Koenen KC, Neylan T, Germine L, Jovanovic T, et al. The AURORA Study: a longitudinal, multimodal library of brain biology and function after traumatic stress exposure. *Mol Psychiatry*. 2020;25:283–96.
43. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–64.
44. Choi SW, O'Reilly PF. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience*. 2019;8:giz082.
45. 1000 Genome Projects Consortium. A global reference for human genetic variation. *Nature*. 2015;526:68–74.
46. Van Rossum G, Drake FL. Python 3 reference manual. Scotts Valley, CA: CreateSpace; 2009.
47. R Core Team. R: a language and environment for statistical computing. 2023. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
48. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. *J Trauma Stress*. 2015;28:489–98.
49. Knol MJ, VanderWeele TJ, Groenwold RHH, Klungel OH, Rovers MM, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. *Eur J Epidemiol*. 2011;26:433–8.
50. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The PhenX Toolkit: get the most from your measures. *Am J Epidemiol*. 2011;174:253–60.
51. Mennis J, Stahler GJ, El Magd SA, Baron DA. How long does it take to complete outpatient substance use disorder treatment? Disparities among Blacks, Hispanics, and Whites in the US. *Addict Behav*. 2019;93:158–65.
52. The TOPPS-II Interstate Cooperative Study Group. Drug treatment completion and post-discharge employment in the TOPPS-II Interstate Cooperative Study. *J Subst Abuse Treat*. 2003;25:9–18.
53. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment*. 2011;18:263–83.
54. Patient-Reported Outcomes Measurement Information System (PROMIS) Health Organization and PROMIS Cooperative Group. *PROMIS depression scoring manual*. HealthMeasures, River Forest, IL; 2015. https://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manuals/_PROMIS_Depression_Scoring_Manual.pdf (accessed 23 May 2024).
55. Acierno R, Kilpatrick DG, Resnick H, Saunders B, De Arellano M, Best C. Assault, PTSD, family substance use, and depression as risk factors for cigarette use in youth: findings from the national survey of adolescents. *J Trauma Stress*. 2000;13:381–96.
56. Davis JP, Diguiseppi G, De Leon J, Prindle J, Sedano A, Rivera D, et al. Understanding pathways between PTSD, homelessness, and substance use among adolescents. *Psychol Addict Behav*. 2019;33:467–76.
57. Gros DF, Flanagan JC, Korte KJ, Mills AC, Brady KT, Back SE. Relations among social support, PTSD symptoms, and substance use in veterans. *Psychol Addict Behav*. 2016;30:764–70.
58. Tripp JC, Jones JL, Back SE, Norman SB. Dealing with complexity and comorbidity: comorbid PTSD and substance use disorders. *Curr Treat Options Psych*. 2019;6:188–97.
59. Wickham H. Tidy data. *J Stat Softw*. 2014;59:1–23.
60. Rubin DB. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, New York, NY; 1987.
61. Keller MC. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol Psychiatry*. 2014;75:18–24.
62. Association AP. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing, Washington DC; 2013.
63. Sullivan PF, Kendler KS. The genetic epidemiology of smoking. *Nicotine Tob Res*. 1999;1:S51–S57.
64. Barr PB, Ksinan A, Su J, Johnson EC, Meyers JL, Wetherill L, et al. Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. *Transl Psychiatry*. 2020;10:196.
65. McKeever VM, Huff ME. A diathesis-stress model of posttraumatic stress disorder: ecological, biological, and residual stress pathways. *Rev Gen Psychol*. 2003;7:237–50.
66. Scheeringa MS. Reexamination of diathesis stress and neurotoxic stress theories: a qualitative review of pre-trauma neurobiology in relation to posttraumatic stress symptoms. *Int J Methods Psychiatr Res*. 2021;30:e1864.
67. Debell F, Fear NT, Head M, Batt-Rawden S, Greenberg N, Wessely S, et al. A systematic review of the comorbidity between PTSD and alcohol misuse. *Soc Psychiatry Psychiatr Epidemiol*. 2014;49:1401–25.
68. Testa M, Livingston JA, Hoffman JH. Does sexual victimization predict subsequent alcohol consumption? A prospective study among a community sample of women. *Addict Behav*. 2007;32:2926–39.
69. Read JP, Colder CR, Merrill JE, Ouimette P, White J, Swartout A. Trauma and posttraumatic stress symptoms predict alcohol and other drug consequence trajectories in the first year of college. *J Consult Clin Psychol*. 2012;80:426–39.
70. Breslau N, Davis GC, Schultz LR. Posttraumatic stress disorder and the incidence of nicotine, alcohol, and other drug disorders in persons who have experienced trauma. *Arch Gen Psychiatry*. 2003;60:289–94.
71. Beseler CL, Aharonovich E, Hasin DS. The enduring influence of drinking motives on alcohol consumption after fateful trauma. *Alcohol Clin Exp Res*. 2011;35:1004–10.
72. Pierce B, Kirsh T, Ferguson AR, Neylan TC, Ma S, Kummerfeld E, et al. Causal discovery replicates symptomatic and functional interrelations of posttraumatic stress across five patient populations. *Front Psychiatry*. 2022;13:1018111.
73. Greenberg JB, Ameringer KJ, Trujillo MA, Sun P, Sussman S, Brightman M, et al. Associations between posttraumatic stress disorder symptom clusters and cigarette smoking. *Psychol Addict Behav*. 2012;26:89–98.
74. Trindade Fortes J, Giordani Cano F, Alcoforado Miranda V, Chung Kang H, Fontenelle LF, Mendlowicz MV, et al. PTSD predicts smoking cessation failure in a trauma-exposed population. *J Dual Diagn*. 2020;16:392–401.
75. Mathew AR, Cook JW, Japuntich SJ, Leventhal AM. Post-traumatic stress disorder symptoms, underlying affective vulnerabilities, and smoking for affect regulation. *Am J Addict*. 2015;24:39–46.
76. Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addict Behav*. 2017;73:133–6.

ACKNOWLEDGEMENTS

The investigators wish to thank the trauma survivors participating in the AURORA Study. Their time and effort during a challenging period of their lives make our efforts to improve recovery for future trauma survivors possible. This project was supported by NIMH under U01MH110925, the US Army MRMC, One Mind, and The Mayday Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funders.

AUTHOR CONTRIBUTIONS

HGD and CAD conceptualized the research question and methodology for this analysis. HGD conducted the formal analyses, generated tables and visualizations, and drafted the original manuscript. CAH, JDT, JLM, SDL, and KCK provided review of the initial analyses and manuscript. SLH, FLB, XA, JSS, TCN, JDC, LTG, SLR, JPH, ABS, PIM Jr., PLH, SS, CWJ, BEP, JLP, MJS, EH, CP, DAP, RCM, RMD, BJON, PS, LDS, SEB, and SEH, SAM, KJR, and KCK contributed to the cohort study design, data collection and investigation, and reviewed and edited the drafted manuscript. SAM, KJR, and KCK also supervised funding acquisition, the cohort data collection, data curation, and overall project administration of the cohort. All authors reviewed and approved the manuscript.

COMPETING INTERESTS

Dr. Garrison-Desany has no competing interests related to this work, though received a contract from Indivior PLC from February to May 2025 for research unrelated to this manuscript. Dr. Hinojosa receives support from the National Institute of Alcohol Abuse and Alcoholism, K99AA031333. Dr. Neylan has received research support from NIH, VA, and Rainwater Charitable Foundation, and consulting income from Jazz Pharmaceuticals. In the last three years Dr Clifford has received research funding from the NSF, NIH and LifeBell AI, and unrestricted donations from AliveCor Inc, Amazon Research, the Center for Discovery, the Gates Foundation, Google, the Gordon and Betty Moore Foundation, MathWorks, Microsoft Research, Nextsense Inc, One Mind Foundation, and the Rett Research Foundation. Dr Clifford has financial interest in AliveCor Inc and Nextsense Inc. He also is the CTO of MindChild Medical with significant stock. These relationships are unconnected to the current work. Dr. Germine receives funding from the National Institute of Mental Health (R01 MH121617) and is on the board of the Many Brains Project. Her family also has equity in Intelrad Medical Systems, Inc. Dr. Rauch reported serving as secretary of the Society of Biological Psychiatry; serving as a board member of Community Psychiatry and Mindpath Health; serving as a board member of National Association of Behavioral Healthcare; serving as secretary and a board member for the Anxiety and Depression Association of America; serving as a board member of the National Network of Depression Centers; receiving royalties from Oxford University Press, American Psychiatric Publishing Inc, and Springer Publishing; and receiving personal fees from the Society of Biological Psychiatry, Community Psychiatry and Mindpath Health, and National Association of Behavioral Healthcare outside the submitted work. Dr. Jones has no competing interests related to this work, though he has been an investigator on studies funded by AstraZeneca, Vapotherm, Abbott, and Ophirex. Dr. Harte has no competing interest related to this work, though in the last three years he has received research funding from Aptinyx and Arbor Medical Innovations, and consulting payments from Memorial Sloan Kettering Cancer Center, Indiana University, The Ohio State University, and Dana Farber Cancer Institute. Dr. McLean has served as a consultant for Walter Reed Army Institute for Research, Arbor Medical Innovations, and BioXcel Therapeutics, Inc. Dr. Ressler has performed scientific consultation for BioRxcel, Bionomics, Acer, and Jazz Pharma; serves on Scientific Advisory Boards for Sage, Boehringer Ingelheim, Senseye, and the Brain Research Foundation, and he has received sponsored research support from Alto Neuroscience. Dr. Koenen's has been a paid scientific consultant for the US

Department of Justice and Covington Burling, LLP over the last three years. She receives royalties from Guilford Press and Oxford University Press.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-025-03593-z>.

Correspondence and requests for materials should be addressed to Henri M. Garrison-Desany.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2025

¹Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ²Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, USA. ³Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ⁴Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. ⁵Stanley Center for Psychiatric Research, Broad Institute, Boston, MA, USA. ⁶Department of Psychiatry and Behavioral Sciences, State University of New York Downstate Medical Center, New York, NY, USA. ⁷Institute for Trauma Recovery, Department of Anesthesiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁸Department of Emergency Medicine, Washington University School of Medicine, St. Louis, MO, USA. ⁹Department of Epidemiology, Brown University, Providence, RI, USA. ¹⁰Department of Emergency Medicine, Brown University, Providence, RI, USA. ¹¹Departments of Psychiatry and Neurology, University of California San Francisco, San Francisco, CA, USA. ¹²Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA, USA. ¹³Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, USA. ¹⁴Institute for Technology in Psychiatry, McLean Hospital, Belmont, MA, USA. ¹⁵The Many Brains Project, Belmont, MA, USA. ¹⁶Department of Psychiatry, Harvard Medical School, Boston, MA, USA. ¹⁷Department of Psychiatry, McLean Hospital, Belmont, MA, USA. ¹⁸Department of Emergency Medicine, University of Massachusetts Chan Medical School, Worcester, MA, USA. ¹⁹Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ²⁰Department of Emergency Medicine, Indiana University School of Medicine, Indianapolis, IN, USA. ²¹Department of Emergency Medicine, University of Florida College of Medicine - Jacksonville, Jacksonville, FL, USA. ²²Department of Emergency Medicine, Cooper Medical School of Rowan University, Camden, NJ, USA. ²³Department of Emergency Medicine, Ohio State University College of Medicine, Columbus, OH, USA. ²⁴Ohio State University College of Nursing, Columbus, OH, USA. ²⁵Department of Surgery, Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA, USA. ²⁶Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ²⁷Department of Surgery, Division of Traumatology, Surgical Critical Care and Emergency Surgery, University of Pennsylvania, Philadelphia, PA, USA. ²⁸Department of Emergency Medicine, Einstein Medical Center, Philadelphia, PA, USA. ²⁹Department of Emergency Medicine, Wayne State University, Ascension St. John Hospital, Detroit, MI, USA. ³⁰Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, USA. ³¹Department of Emergency Medicine, Harvard Medical School, Boston, MA, USA. ³²Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, USA. ³³Department of Emergency Medicine, Trinity Health-Ann Arbor, Ypsilanti, MI, USA. ³⁴Department of Emergency Medicine, Wayne State University, Detroit Receiving Hospital, Detroit, MI, USA. ³⁵Department of Emergency Medicine, McGovern Medical School at UTHealth, Houston, TX, USA. ³⁶Department of Psychological Sciences, University of Missouri - St. Louis, St. Louis, MO, USA. ³⁷Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, MI, USA. ³⁸Department of Internal Medicine-Rheumatology, University of Michigan Medical School, Ann Arbor, MI, USA. ³⁹Department of Emergency Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁴⁰Institute for Trauma Recovery, Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁴¹Division of Depression and Anxiety, McLean Hospital, Belmont, MA, USA. ⁴²Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. [✉]email: mgarris7@jhu.edu