

1 **Disentangling the roles of trauma and genetics in psychiatric disorders using an Electronic**
2 **Health Records-based approach**

3

4 Shelby Marchese^{1,3}, Winston Cuddleston^{1,3}, Carina Seah^{1,3}, Jessica Johnson^{1,3}, Laura M.
5 Huckins¹⁻⁵

6

7 ¹ Pamela Sklar Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai,

8 New York, NY 10029, USA

9 ² Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029,

10 USA

11 ³ Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New

12 York, NY 10029, USA

13 ⁴ Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai,

14 New York, NY 10029, USA

15 ⁵ Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai,

16 New York, NY 10029, USA

17

18 *Correspondence should be addressed to Dr. Laura Huckins (email: laura.huckins@mssm.edu;

19 telephone: 212-659-1613; mailing address: Icahn School of Medicine at Mount Sinai, 1 Gustave

20 Levy Place, NYC, NY, 10029, USA)

21

22

23 **Abstract**

24 Posttraumatic stress disorder (PTSD) requires an exposure to trauma for diagnosis by the DSM-
25 V. Despite this, there is no documented linear relationship between degree of trauma and
26 severity/chronicity of PTSD.

27 To determine whether traumatic and stressful life events (TSLEs) collected from Electronic
28 Health Records (EHR) interact with PTSD genetics to better define individual risk of developing
29 PTSD. We collected trauma information from patient records in the Mount Sinai BioMeTM
30 biobank population-based cohort and tested for associations with PTSD. We generated a TSLE
31 risk score (TRS), tested its association with PTSD, and for interactions with a polygenic risk
32 score (PRS) of PTSD and a GWAS of PTSD using our biobank population.

33 We used the Mount Sinai BioMeTM biobank patient population of 31,704 individuals with
34 matched genotype and EHR data, which has been enrolling patients since 2006. Patient
35 enrollment in BioMeTM is unrestricted, resulting in high diversity. Our study includes 1,990
36 individuals with PTSD and 28,478 individuals without PTSD from the Mount Sinai BioMeTM
37 biobank.

38 A total of 1,990 individuals with PTSD and 28,478 controls were analyzed (average age 42-78
39 years, 58.7% women). We identified a total of 22 EHR-derived TSLEs that were associated with
40 PTSD ($\beta > 0.029$, $p < 1.61 \times 10^{-3}$). TSLEs interacted with each other to increase the risk of
41 developing PTSD, with the most significant interaction between being female (as a proxy for
42 experiencing sexism) and being HIV+ ($\beta = 0.013$, $p = 8.9 \times 10^{-11}$). PRS of PTSD and lead SNPS
43 from GWAS interacted with TSLEs and our TRS to increase PTSD risk. In addition to TRS
44 interactions, we found significant interactions between PTSD PRS and domestic violence, and
45 sexual assault in European Americans ($\beta > 207.74$, $p < 1.80 \times 10^{-3}$). rs113282988 and rs189796944
46 variants reached genome-wide significance in interactions with TRS ($\beta > 0.056$, $p < 9.04 \times 10^{-9}$),
47 and rs189796944 in an interaction with Physical Survival TSLEs ($\beta > 0.127$, $p < 4 \times 10^{-8}$).

48 In conclusion, quantification of trauma type, severity, and magnitude—alone and in concert with
49 genetics—provides better prediction of PTSD risk than PRS alone. Understanding who is at risk
50 of developing PTSD allows for timely clinical intervention and possible prevention.

51 **Introduction**

52 Posttraumatic Stress Disorder (PTSD) is characterized by symptoms of re-experiencing,
53 avoidance, emotional distress, and heightened physiological arousal following exposure to a
54 traumatic event. 65-89% of all individuals will face a significant traumatic or stressful life event
55 (TSLE) in their lifetime¹⁻³, but only 10-30% will develop PTSD^{4,5}. Understanding the relative
56 roles of trauma and genetics in PTSD diagnosis will be key to identifying who might develop
57 PTSD following trauma. Therefore, in this study we combine measures of TSLEs with genetic
58 measures such as polygenic risk scores (PRS) and genome-wide association study (GWAS)
59 results using a large hospital-based biobank.

60

61 The genetic etiology of PTSD varies according to gender⁶⁻⁸, and index trauma type⁹⁻¹¹;
62 consequently, studies that examine relatively homogeneous cohorts – i.e. individuals exposed to
63 only one specific type of trauma¹²⁻¹⁸ – have identified genome-wide significant associations that
64 have not yet been identified using meta-analytic approaches across trauma types. Reducing
65 cohort heterogeneity by characterizing trauma is imperative to advancing PTSD genetic
66 discovery. Characterizing individuals according to a single index trauma may help, but could fail
67 to fully capture trauma histories among cohorts. Single index trauma characterization cannot
68 account for degree or chronicity, and may fail to identify truly causal traumas within the cohort.
69 Therefore, in this study we combine various types of Electronic Health Record (EHR)-derived
70 TSLEs to construct individual-level trauma scores to capture chronicity and degree of trauma
71 exposure.

72

73 The benefits of using EHR data are numerous. First, data is readily available and does not have
74 to be collected, reducing the time it takes to conduct a study¹⁹. Second, EHR encompasses a

75 more diverse range of individuals, potentially including all patients entering the hospital
76 system²⁰. Third, the range of TSLEs available may allow us to identify unexpected relationships
77 with PTSD or other psychiatric disorders. For example, interactions between polygenic risk score
78 (PRS) and childhood trauma (CT) explain more variability in Major Depressive Disorder (MDD)
79 than PRS alone^{37,21,22}, and CT has been linked to bipolar disorder^{23,24}, schizophrenia²⁴, and
80 cognitive decline²⁵.

81
82 In this study, we used EHR data from the Mount Sinai BioMeTM biobank (BioMeTM) to identify
83 22 TSLEs, and tested for associations with PTSD. We found significant associations between
84 individual TSLEs and groups of TSLEs with PTSD. Additionally, we found that multiple TSLEs
85 interact to increase disease risk. We also discovered significant interactions between TSLEs and
86 PTSD-PRS, suggesting a that trauma interacts with common variants to increase disease risk. To
87 further support this finding, we performed a PTSD-GWAS and looked at the interaction between
88 significant SNPs and TSLEs, which increased the variance explained in PTSD. We hope that
89 these findings will allow better understanding of the complexity of trauma in PTSD and improve
90 patient intake, diagnostics, and future care.

91

92

93 **Methods**

94 **Participants from the Mount Sinai BioMeTM biobank**

95 Our study includes 31,704 individuals from BioMeTM, with matched genotype²⁶ and EHR data.
96 BioMeTM is highly diverse (self-reported ancestry: HA, 35%; EA, 34%; AA, 25%; Other 6%),
97 with an average age of 59.84 (SE=17.85), and an average EHR records length of 10.4 (SE=4.2)
98 years. A total of 58.7% of the sample self-identify as women; 39.2% as men; 2.1% as non-
99 binary.

100

101 **Defining TSLEs**

102 Since our sample has not been collected for a particular study of trauma, but rather is an existing
103 patient biobank, trauma phenotypes may not have been equally assessed across the sample. For
104 example, patients already seeing a psychiatrist are more likely to be asked about trauma.
105 Therefore, we sought to identify TSLEs that might be more generally captured in the EHR, such
106 as serious illness and self-actualization traumas, and additionally identified positive, negative,
107 and null trauma controls. We selected positive controls, including codes that may be proxies or
108 pre-cursors to PTSD diagnoses (e.g. grief); negative controls that we expect to be protective
109 (e.g., marriage), and null controls that we expect to be unassociated with PTSD or TSLEs.

110

111 We included only TSLEs occurring in at least 10 patients. We divided TSLEs into five literature-
112 defined groups²⁷ (**Table 1**): personal (e.g. domestic violence), exposure to systematic
113 discrimination (e.g., racism, homophobia, inferred by proxies), self-actualization (e.g.
114 homelessness), physical survival (e.g., serious illness) and interdependence (affecting a direct
115 relative).

116 **Defining PTSD**

117 We used a broad case-definition of PTSD which included all International Classification of
118 Disease (ICD)-10 codes (≥ 1 count) under the F43 group classification. Controls included all
119 individuals with zero counts under the F43 group classification.

120

121 **Trauma contribution to PTSD**

122 We tested for associations between each TSLE and ICD-code defined PTSD using a linear
123 regression model, corrected for age. We applied a Bonferroni correction to establish significance,
124 correcting for the number of TSEs and controls tested.

125

126 **Quantifying trauma by individual**

127 In order to assess the ‘amount’ of TSLEs experienced by each individual within our biobank, we
128 created three trauma scores (TRS): first, a simple aggregate trauma score (TRS_{SUM}); a score
129 weighted by frequency of each TSLE for each individual (TRS_{FREQ}); and a score derived from a
130 joint elastic net model (TRS_{EN}). To create our TRS_{EN} , we applied a 10-fold cross-validation
131 approach using the “cvglmnet” function in *glmnet* elastic net regression package in R²⁸. We
132 define each of our trauma scores as follows (Eq. 1):

133

$$TRS_{SUM,i} = \sum TSLE_{i,j}$$

$$TRS_{FREQ} = \sum TSLE_{i,j} * freq(TSLE_{i,j})$$

$$TRS_{EN} = \sum TSLE_{i,j} * \beta_j$$

134

135

Where:

$$TSLE_{i,j} = \begin{cases} 1 & \text{if individual } i \text{ has experienced } TSLE_j \\ 0 & \text{otherwise} \end{cases}$$

$freq(TSLE_{i,j}) = \text{frequency of individual } i \text{ experiencing } TSLE_j$

$\beta_j = \text{elastic - net derived effect size for } TSLE_j$

136

137 *Equation 1. Trauma risk scores derived from Electronic Health Records data.*

138

139 **PRS were calculated for PTSD ancestries**

140 We tested the PTSD PRS of specific ancestries with $N > 100$ PTSD cases; these were AA-PTSD

141 ($R^2=0.02$, $P_t=2.5 \times 10^{-4}$, $p=0.011$), EA-PTSD ($R^2=0.07$, $P_t=0.028$, $p=0.142$) and HA-PTSD

142 ($R^2=0.049$, $P_t=0.32$, $p=0.028$). PRS were derived from the largest Psychiatric Genetics

143 Consortium (PGC) PTSD GWAS⁶ separately for each ancestry. We fit the best model for each

144 ancestry using PRSice-2²⁹. Variance explained was calculated correcting for age.

145

146 **PRS were tested for interactions with trauma**

147 We tested for interactions between trauma and PTSD PRS using a linear model, correcting for

148 age (Eq. 2).

$$PTSD \sim \text{Polygenic Risk Score} * TSLE + \text{Age}$$

149

150 *Equation 2. Linear regression to determine which TSLEs interact with PRS for PTSD.*

151

152 **GWAS meta-analysis of all three ancestries**

153 We performed three ancestry-specific PTSD GWAS, in AA, HA, and EA samples within
154 BioMeTM; we excluded other ancestries with <100 cases. We defined PTSD using a broad
155 definition in line with PGC-PTSD GWAS³⁰, resulting in a total sample size across all three
156 cohorts of 1,990 PTSD patients and 28,478 controls. SNPS that had a MAF of <1%, Hardy-
157 Weinberg equilibrium $P > 0.001$ or call rate <98% were excluded. GWAS were calculated using
158 the first 10 principal components and age as covariates. Meta-analysis across all ancestry groups
159 was performed using an inverse-variance-weighted fixed effects model with METAL³¹.

160

161 **Finemapping of GWAS meta-analysis results**

162 We performed finemapping on all GWAS loci including at least one SNP reaching $p < 1 \times 10^{-5}$ in
163 our meta-analysis to identify putatively causal SNPs. Finemapping was performed using
164 PolyFun³² in combination with FINEMAP³³. Lead SNPs were analyzed for causal variants
165 contained in the locus 500kb upstream and downstream from start site. Casual variants were
166 selected based on a Posterior Inclusion Probability (PIP) that constitutes a total (PIP=1.0) based
167 on the top two variants in the locus (PIP>0.40 for all variants).

168

169

170 **Results**

171 **Identification of traumatic and stressful life events in the Mount Sinai BioMeTM biobank**

172 We identified 23 traumatic and stressful life events (TSLEs) occurring in >10 individuals within
173 BioMeTM (**Table 1A**). TSLEs were split into 5 categories: personal, exposure to systematic
174 discrimination, self-actualization, physical survival, and interdependence²⁷. In addition, we
175 selected nine positive negative and null controls.

176
177 22/23 TSLEs were significantly associated with PTSD ($p < 1.61 \times 10^{-3}$). The top three associations
178 included surgery ($\beta = 0.069$, $p = 2.95 \times 10^{-120}$), HIV+ ($\beta = 0.145$, $p = 1.37 \times 10^{-103}$), and female sex
179 ($\beta = 0.033$, $p = 6.80 \times 10^{-31}$; **Table 1A**).

180
181 Using a linear regression model weighted by an individual's frequency of exposure to a trauma,
182 we identified one additional TSLE that was significantly associated with PTSD: heart attack
183 ($\beta = 0.012$, $p = 8.47 \times 10^{-4}$; **Supplementary Table 1**). Otherwise, TSLE significance and effect size
184 in association with PTSD was similar between the two models.

185
186 In line with our initial hypotheses, we identified positive associations between PTSD and all four
187 positive controls, and negative associations between PTSD and marriage. Getting a flu vaccine
188 (null control) was not significantly associated with PTSD, as expected. Most significantly,
189 psychotherapy was associated with an increased risk of PTSD ($\beta = 0.504$, $p = 2.56 \times 10^{-75}$), while
190 marriage was most significantly protective against PTSD ($\beta = -0.027$, $p = 3.99 \times 10^{-18}$; **Table 1A**).
191 The other negative/null controls were small in effect size, but significantly positively associated
192 with PTSD (**Table 1A**).

193

194 Next, we hypothesized that specific combinations of TSLEs may represent a greater risk for
195 PTSD. Therefore, we tested for interaction effects between all 21 TSLEs, and identified nine
196 significant interactions. The most significant was being female: HIV ($\beta=0.013$, $p=8.9 \times 10^{-11}$);
197 i.e., women who are infected with (or test positive for) HIV are at greater risk of PTSD than men
198 infected with HIV (**Table 1B**). We found that pairs of TSLEs with significant interactions tended
199 to span categories, rather than coming from the same category (**Figure 1A**).

200

201 **TSLE risk score measures the association between disease risk and TSLEs**

202 We tested whether aggregate TSLE exposure increases risk for PTSD beyond the risk posed by
203 single TSLEs and, if so, whether the number of TSLEs, frequency of each TSLEs, or specific
204 combinations of TSLEs increase the risk further. To do so, we created three trauma risk scores: a
205 simple aggregate trauma score (TRS_{SUM}); a score weighted by frequency of each trauma for each
206 individual (TRS_{FREQ}); and a score derived from a joint elastic net model (TRS_{EN}).

207

208 The total number of TSLEs experienced (TRS_{SUM}) was significantly correlated with PTSD
209 diagnosis ($p=2.57 \times 10^{-229}$), explaining ~3.3% of phenotypic variance within our sample. TRS_{SUM}
210 was significantly higher among PTSD cases compared to controls (TRS_{SUM} average = 2.534 [SD
211 =1.15] vs. 1.718 [SD=1.07]; $p=3.353 \times 10^{-173}$). Similarly, the frequency of each TSLE (TRS_{FREQ})
212 was significantly associated with PTSD ($p=2.56 \times 10^{-270}$), explaining ~4% of the phenotypic
213 variance, and was significantly higher in cases compared to controls (TRS_{FREQ} average = 3.707
214 [SE=1.79] vs. 2.415 [SE=1.56]; **Figure 1C**).

215

216 Finally, TRS_{EN} explained 5.44% of phenotypic variance in PTSD ($p=1.0 \times 10^{-300}$), substantially
217 more than either of the other scores. Like the previous TRS, TRS_{EN} was significantly higher
218 among cases compared to controls (average = 1.04 (SE=0.673) vs. 0.562 (SE=0.461); $p < 1.4 \times 10^{-72}$).
219 No single category of TSLEs drives this effect; in total, our TRS_{EN} includes 17 TSLEs
220 (**Supplementary Table 2**), including TSLEs from each of our five categories. Constructing
221 TRS_{EN} using each category individually also produces significantly different TRS_{EN} scores
222 between cases and controls for every group except self-actualization TSLEs (**Supplementary**
223 **Figure 2**).

224

225 **Contribution of TSLEs to the prevalence and severity of other psychiatric disorders**

226 We performed a phenome-wide association study (PheWAS) to identify other psychiatric traits
227 associated with TRS_{EN} (**Supplementary Figure 3**). We identified 37 psychiatric and substance-
228 use traits, and 25 neurological traits significantly associated with TRS_{EN} at Bonferroni
229 significance (**Table 2**): including anxiety disorders (PheCode 300.1, OR=1.570, $p=8.91 \times 10^{-305}$),
230 adjustment reaction (PheCode 304, OR=2.282, $p=2.46 \times 10^{-206}$), and PTSD (PheCode=300.9,
231 OR=1.888, $p=4.99 \times 10^{-107}$). Similarly, we identified 27 psychiatric and substance-use traits, and
232 25 neurological traits significantly associated with the TRS_{FREQ} at Bonferroni significance
233 (**Supplementary Table 3**). There were 16 disorders that were differentially associated with
234 TRS_{FREQ} and TRS_{EN} , including pain (PheCode 338, OR=1.492, $p=6.25 \times 10^{-180}$), lack of
235 coordination (PheCode 350.3, OR=1.27, $p=1.43 \times 10^{-10}$) and meningitis (PheCode=320,
236 OR=1.327, $p=1.29 \times 10^{-8}$) for TRS_{FREQ} and memory loss (PheCode 292.3, OR=1.489, $p=5.49 \times 10^{-61}$),
237 dementia (PheCode 290.1, OR=1.299, $p=7.98 \times 10^{-17}$) and mild cognitive impairment
238 (PheCode 292.2, OR=1.428, $p=3.71 \times 10^{-13}$) for TRS_{EN} .

239 We also performed a PheWAS on each of our TSLE categories to measure their phenotypic
240 associations (**Table 2**). Except for the negative controls, each category was significantly
241 associated with psychiatric disorders. For systematic discrimination-TSLEs and positive
242 controls, the most highly-associated psychiatric disorder was anxiety disorders (PheCode=300,
243 $OR=2.335$, $p=1.87 \times 10^{-58}$; $OR=3.857$, $p=9.72 \times 10^{-294}$). Interdependence- and physical survival-
244 TSLEs were most highly associated with mood disorders (PheCode=296, $\beta=0.398$, $p=1.81 \times 10^{-54}$;
245 $OR=1.898$, $p=1.42 \times 10^{-222}$). Personal-TSLEs were most highly associated with PTSD
246 (PheCode=300.9, $OR=3.013$, $p=3.41 \times 10^{-12}$), and self-actualization-TSLEs were most highly
247 associated with MDD (PheCode=296.22, $OR=3.456$, $p=1.61 \times 10^{-20}$; **Supplementary Figure 4**).

248

249 **GWAS meta-analysis reveals PTSD SNPs that interact with TSLEs**

250 A key motivator of our study was to investigate whether including trauma in our studies might
251 more deeply elucidate the genetic architecture of PTSD¹². Therefore, we conducted a GWAS
252 within our sample, following PGC broad PTSD diagnostic definitions⁷, and refined our
253 associations using FINEMAP. As expected given our modest sample size, we did not identify
254 any genome-wide significant associations (**Supplementary Figure 5**).

255

256 Based on finemapping results, we selected 18 SNPs ($p < 1 \times 10^{-5}$), and tested for interaction effects
257 with individual TSLEs and TRS. We found six Bonferroni significant SNP x TSLE or SNP x
258 TRS_{EN} interactions ($p < 7.51 \times 10^{-5}$; **Table 3**), and 10 Bonferroni significant SNP x $TSLE_{FREQ}$ or
259 SNP x TRS_{FREQ} interactions ($p < 1.33 \times 10^{-4}$). Of these interactions, five increased in significance
260 compared to the GWAS and reached genome-wide significance: rs189796944 x TRS_{EN} and

261 $TRSFREQ$ ($\beta > 0.060$, $p < 9.04 \times 10^{-9}$), $rs189796944 \times PS-TRSEN$ and $PS-TRSFREQ$ ($\beta > 0.127$, $p < 4 \times 10^{-8}$), and $rs113282988 \times TRSFREQ$ ($\beta > 0.056$, $p < 3.28 \times 10^{-9}$; **Table 3**).

263

264 **PRS elucidate more GxE interactions with PTSD**

265 Finally, we tested whether including TSLEs and TRS in concert with PRS might help to clarify
266 individual-level risk for PTSD. We tested whether PRS significantly predicted PTSD alone and
267 in concert with TSLEs and TRS, in three ancestry-specific groups within our biobank (European,
268 EA; African-American, AA; Hispanic American, HA).

269

270 First, we compared the amount of variance explained when including (i) PRS alone (ii) TRS
271 alone and (iii) an interaction between PRS and TRS (**Supplementary Table 5**). For all
272 ancestries, we found a significant increase in variance explained when including the interaction
273 term, compared to PRS alone ($p < 2.6 \times 10^{-102}$). However, in all circumstances, TRS alone
274 explained the most variance ($R^2 = 3.3-5.4\%$, $p < 2.56 \times 10^{-270}$).

275

276 Next, we considered more specific associations between PRS and TSLEs and TRS. In our
277 European (EA) cohort, we found significant interactions between PTSD-PRS and personal-
278 TSLEs ($TRSEN$; $\beta = 978.55$, $p = 4.71 \times 10^{-4}$), self-actualization-TSLEs ($TRSEN$; $\beta = 1045.52$,
279 $p = 3.66 \times 10^{-3}$), and positive controls (**Table 4**). For individual TSLEs in EA, we found
280 interactions between PTSD-PRS and domestic violence ($\beta = 591.53$, $p = 1.07 \times 10^{-3}$), and between
281 PTSD-PRS and sexual assault ($\beta = 207.74$, $p = 1.80 \times 10^{-3}$). We found a significant interaction with
282 frequency of personal-TSLEs and EA-PTSD-PRS, both overall ($\beta = 264.31$, $p = 4.09 \times 10^{-3}$) and for
283 4 specific TSLEs: psychotherapy, heart attack, work stress, and sexual assault ($\beta > 207.43$,

284 $p < 6.58 \times 10^{-4}$; **Table 4**). We also observed a significant interaction between TRS_{EN} and PTSD-
285 PRS for in our African-American cohort ($\beta = 14.24$, $p = 2.71 \times 10^{-3}$).

286
287 Finally, we considered whether TRS and PRS applied in concert might provide more
288 diagnostically useful risk-stratification. To do so, we compared the proportion of cases and
289 controls falling in the top decile to (i) the bottom decile; and (ii) the bottom 9 deciles, using (i)
290 PRS alone; (ii) TRS alone; (iii) PRS and TRS in concert. Across all ancestries, individuals in the
291 top decile of TRS risk were significantly more likely to be PTSD cases, compared to the bottom
292 decile ($OR > 7.87$, $p < 1.37 \times 10^{-36}$; **S. Table 5**), and the bottom nine deciles ($OR > 3.75$, $p < 8.41 \times 10^{-43}$).
293 Individuals in **both** the top TRS **and** PRS deciles were more significantly enriched for cases
294 in AA and HA populations ($OR > 8.50$, $p < 2.62 \times 10^{-8}$). Although numbers are small, the same
295 pattern holds when considering the top quintile (i.e., top 20%; **S. Table 5B**)

296

297

298

299

300

301

302 **Discussion**

303 In addition to DSM-5 recognized PTSD-associated traumas (e.g. sexual assault, domestic
304 violence, grief), we also discovered a large number of atypical TSLEs associated with PTSD,
305 such as systematic discrimination measures (e.g. racism, sexism, homophobia), stress from work,
306 finances, and other physical health stressors such as cancer or experiencing a heart attack. We
307 examined whether these interactions occur across or within TSLE categories to address whether
308 two similar TSLEs contribute more risk for PTSD than two independent TSLEs. We found that
309 while TSLEs do interact, the majority of interactions are between different TSLE categories,
310 rather than within the same category. Interaction of these TSLEs leads to a higher risk of
311 developing PTSD, and thus led us to investigate the risk and implications of experiencing
312 multiple traumatic events.

313
314 To expand on the idea that multiple TSLEs lead to a higher risk and more severe PTSD
315 symptoms, we constructed an aggregate trauma risk score (TRS) of TSLEs for an individual and
316 found that using elastic net-derived effect sizes predicts more phenotypic variance in PTSD than
317 an unweighted model, exposure-counts weighted-model, or PTSD-PRS. On average, individuals
318 with PTSD experience a higher number and frequency of TSLEs.

319
320 The association between trauma exposure and risk of developing a psychiatric disorder has been
321 demonstrated in psychiatric disorders other than PTSD, such as MDD^{22,34–37}, substance abuse^{38–}
322 ⁴⁰, and suicidality³⁸. However, the influence of trauma-specific factors (type, chronicity, etc.) on
323 disease risk has not been well studied for the majority of psychiatric disorders. Our PheWAS
324 analysis suggests that trauma is significantly associated with many psychiatric disorders. As
325 such, psychiatric disorder studies may benefit from the incorporation of trauma information into

326 the study design. To accomplish this, trauma screening questionnaires should be implemented as
327 a standard tool in all psychiatric disorder studies.

328
329 Two loci reached genome-wide significance in interactions with our TRS: rs113282988 (an
330 intronic variant in *ITGAE*) and rs189796944 (an intronic variant in *DPYSL2*). *DPYSL2*, also
331 known as *CRMP2*, is involved in facilitating neuron guidance, growth and polarity⁴¹⁻⁴³. It is also
332 implicated in multiple neurological disorders⁴⁴ including Alzheimer's disease^{45,46},
333 schizophrenia⁴⁷⁻⁵², and alcohol-dependence⁵³. *ITGAE* canonically mediates the adhesion of intra-
334 epithelial T-lymphocytes to epithelial cells and is involved in intra-tumoral immunity⁵⁴⁻⁵⁶, but it
335 has also been implicated in Parkinson's disease⁵⁷ and Attention-deficit/hyperactivity disorder
336 (ADHD)⁵⁸⁻⁶⁰.

337
338 We found that our TRS significantly improves the variance explained in PTSD over PRS. We
339 see significant interaction between PRS and TRS ($p < 1.93 \times 10^{-49}$), suggesting that common
340 variants, in concert with TRS, increase PTSD risk. In addition to TRS interactions, we found
341 interactions between PRS and domestic violence, and sexual assault in EA. It has been well
342 documented that these TSLEs can individually lead to PTSD⁶¹⁻⁶⁵, but this is the first study to
343 specifically test whether the multiplicity of experiencing these TSLEs can lead to a higher risk of
344 developing PTSD. Additionally, we found that individuals who into the top decile (or quintile) of
345 both TRS and PRS were significantly enriched for PTSD (OR 4.18-10.05). In both African
346 American and Hispanic American cohorts, individuals in the top TRS and PRS decile were more
347 likely to have PTSD than individuals in the top TRS decile alone; however, this was not the case
348 for Europeans. We hypothesize that this may be due to lower overall levels of trauma

349 experienced in this population; it is possible that individuals who do report trauma are more
350 likely to do so because of a PTSD diagnosis, increasing the predictive accuracy of the TRS. In
351 line with this, the OR for PTSD cases in the top vs. bottom TRS decile was substantially larger
352 for European Americans (OR=12.68) than for African (OR=9.62) or Hispanic Americans
353 (OR=7.87).

354
355 There are three potential reasons why our TRS significantly outperforms the PTSD-PRS: (i) the
356 model for TRS is potentially overfit, (ii) the relationship between TRS and PRS in controls,
357 and/or (iii) our TRS itself might be more heritable than PTSD. Our TRS was designed using
358 trauma exposures within our own cohort, and while we used elastic net regression with ten-fold
359 cross validation to attempt to minimize statistical bias, the model may potentially be overfit.
360 Additionally, at lower PRS, the risk of developing PTSD is two-fold lower in individuals with a
361 high TRS compared to a low TRS (**Supplementary Figure 6**). This suggests a potential genetic
362 resilience in individuals with a low PRS. In controls exposed to multiple traumas, individuals
363 with a low PRS are more protected from developing PTSD than those with a high PRS. This
364 interaction may also be moderated by gene expression or epigenetics. Last, we cannot rule out
365 the possibility that our TRS itself is heritable. In a recent PTSD genome-wide by environment
366 interaction study (GWEIS) study, authors found that their unweighted, quantitative trauma risk
367 score had large genetic overlap with PTSD and reduced PTSD heritability by 31% when adjusted
368 for trauma⁶⁶. Similarly, combat^{30,67} and other traumas⁶⁸ have been reported to be heritable.

369

370 One significant caveat of this study is the lack of adequate sample size for our GWAS and PRS.

371 We did not reach genome-wide significance in the meta-analysis, and our PRS were very low
372 compared to efforts from the PGC, which found SNP-heritability to be 5-20%, varying by sex⁶⁹.

373

374 The diversity and breadth of BioMeTM gave us a unique opportunity to understand the role of
375 trauma in a clinical population. However, EHR-derived data has some limitations: records are
376 easily obtainable and searchable, but they are not curated for the purposes of this study. We are
377 limited in the number of TSLEs that we can test due to recorded availability and low volume of
378 cases. We chose to exclude any TSLE that had <10 cases, due to the potential to highly skew
379 interactions when limited to a small number of individuals. One example of a TSLE that was
380 excluded was childhood trauma (CT), since so few codes are reported (<10), likely indicating
381 under-reporting rather than an absence of the exposure. To include CT would therefore (1) not
382 capture all the individuals with unreported CT and (2) skew the results of interactions with CT to
383 extreme cases only.

384

385 Accounting for trauma increases the amount of variance explained in PTSD, compared to
386 considering genetics alone. Specifically, considering aggregate scores of trauma—weighted
387 according to machine-learning approaches—accounts for a >5-fold increase in the amount of
388 variance explained in PTSD. This suggests that trauma is a critical component in understanding
389 one's risk of developing PTSD, and a thorough trauma history is necessary to help diagnose and
390 treat patients with PTSD. By examining the interactions between trauma and genetics, we can
391 better understand the etiology of PTSD and design better treatments and prevention methods.

392

394 **Conflicts of Interest**

395 The authors declare no conflicts of interest.

396

397

398

399 **References**

400

401 1. Koenen, K. C. *et al.* Posttraumatic stress disorder in the World Mental Health Surveys.

402 *Psychological Medicine* **47**, 2260–2274 (2017).

403 2. Stein, M. B., McQuaid, J. R., Pedrelli, P., Lenox, R. & McCahill, M. E. Posttraumatic stress

404 disorder in the primary care medical setting. *Gen Hosp Psychiatry* **22**, 261–269 (2000).

405 3. Kilpatrick, D. G. *et al.* National Estimates of Exposure to Traumatic Events and PTSD

406 Prevalence Using DSM-IV and DSM-5 Criteria. *J Trauma Stress* **26**, 537–547 (2013).

407 4. Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M. & Nelson, C. B. Posttraumatic stress

408 disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* **52**, 1048–1060 (1995).

409 5. Hoppen, T. H. & Morina, N. The prevalence of PTSD and major depression in the global

410 population of adult war survivors: a meta-analytically informed estimate in absolute

411 numbers. *Eur J Psychotraumatol* **10**, 1578637.

412 6. Duncan, L. E. *et al.* Largest GWAS of PTSD (N=20□070) yields genetic overlap with

413 schizophrenia and sex differences in heritability. *Mol. Psychiatry* **23**, 666–673 (2018).

414 7. Nievergelt, C. M. *et al.* International meta-analysis of PTSD genome-wide association

415 studies identifies sex- and ancestry-specific genetic risk loci. *Nat Commun* **10**, 4558 (2019).

416 8. Kim, G. S. *et al.* Methylomic profiles reveal sex-specific shifts in leukocyte composition

417 associated with post-traumatic stress disorder. *bioRxiv* 594127 (2019) doi:10.1101/594127.

418 9. Agorastos, A. *et al.* The cumulative effect of different childhood trauma types on self-

419 reported symptoms of adult male depression and PTSD, substance abuse and health-related

420 quality of life in a large active-duty military cohort. *J Psychiatr Res* **58**, 46–54 (2014).

421 10. Huckins, L. M. *et al.* Analysis of Genetically Regulated Gene Expression identifies a trauma

422 type specific PTSD gene, SNRNP35. *bioRxiv* 581124 (2019) doi:10.1101/581124.

- 423 11. Kelley, L. P., Weathers, F. W., McDevitt-Murphy, M. E., Eakin, D. E. & Flood, A. M. A
424 comparison of PTSD symptom patterns in three types of civilian trauma. *Journal of*
425 *Traumatic Stress* **22**, 227–235 (2009).
- 426 12. Marchese, S. *et al.* Altered gene expression and PTSD symptom dimensions in World Trade
427 Center responders. *Mol Psychiatry* 1–22 (2022) doi:10.1038/s41380-022-01457-2.
- 428 13. Huckins, L. M. *et al.* Polygenic regulation of PTSD severity and outcomes among World
429 Trade Center responders. *medRxiv* 2020.12.06.20244772 (2020)
430 doi:10.1101/2020.12.06.20244772.
- 431 14. Dunn, E. C. *et al.* Interaction between genetic variants and exposure to Hurricane Katrina on
432 post-traumatic stress and post-traumatic growth: A prospective analysis of low income
433 adults. *Journal of Affective Disorders* **152–154**, 243–249 (2014).
- 434 15. Mehta, D. *et al.* Childhood maltreatment is associated with distinct genomic and epigenetic
435 profiles in posttraumatic stress disorder. *Proc Natl Acad Sci U S A* **110**, 8302–8307 (2013).
- 436 16. Akosile, W. *et al.* NLRP3 is associated with coronary artery disease in Vietnam veterans.
437 *Gene* **725**, 144163 (2020).
- 438 17. Stein, M. B. *et al.* Genomewide Association Studies of Posttraumatic Stress Disorder in Two
439 Cohorts of US Army Soldiers. *JAMA Psychiatry* **73**, 695–704 (2016).
- 440 18. Gelernter, J. *et al.* Genome-wide Association Study of Posttraumatic Stress Disorder (PTSD)
441 Re-Experiencing Symptoms in >165,000 US Veterans. *Nat Neurosci* **22**, 1394–1401 (2019).
- 442 19. Ross, M. K., Wei, W. & Ohno-Machado, L. “Big Data” and the Electronic Health Record.
443 *Yearb Med Inform* **9**, 97–104 (2014).

- 444 20. Hindorff, L. A., Bonham, V. L. & Ohno-Machado, L. Enhancing diversity to reduce health
445 information disparities and build an evidence base for genomic medicine. *Personalized*
446 *Medicine* **15**, 403–412 (2018).
- 447 21. Mullins, N. *et al.* Polygenic interactions with environmental adversity in the aetiology of
448 major depressive disorder. *Psychological Medicine* **46**, 759–770 (2016).
- 449 22. Peyrot, W. J. *et al.* Does Childhood Trauma Moderate Polygenic Risk for Depression? A
450 Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biological*
451 *Psychiatry* **84**, 138–147 (2018).
- 452 23. Moraes*, J. B. *et al.* Elevated C-reactive Protein Levels in Women with Bipolar Disorder
453 may be Explained by a History of Childhood Trauma, Especially Sexual Abuse, Body Mass
454 Index and Age. *CNS & Neurological Disorders - Drug Targets*
455 <http://www.eurekaselect.com/151443/article> (2017).
- 456 24. Aas, M. *et al.* The role of childhood trauma in bipolar disorders. *International Journal of*
457 *Bipolar Disorders* **4**, 2 (2016).
- 458 25. Marchese, S. *et al.* FKBP5 Genotype, Cognition, and Posttraumatic Stress Disorder Severity:
459 Evidence from A Study of 9/11 World Trade Center Rescue and Recovery Workers.
460 *Manuscript in preparation.* (2022).
- 461 26. Zheutlin, A. B. *et al.* Penetrance and Pleiotropy of Polygenic Risk Scores for Schizophrenia
462 in 106,160 Patients Across Four Health Care Systems. *AJP* **176**, 846–855 (2019).
- 463 27. Kira, I. A. *et al.* Measuring Cumulative Trauma Dose, Types, and Profiles Using a
464 Development-Based Taxonomy of Traumas. [Miscellaneous Article]. *Traumatology* **14**, 62–
465 87 (2008).

- 466 28. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear
467 Models via Coordinate Descent. *Journal of Statistical Software* **33**, 1–22 (2010).
- 468 29. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale
469 data. *GigaScience* **8**, giz082 (2019).
- 470 30. Nievergelt, C. M. *et al.* Genomic predictors of combat stress vulnerability and resilience in
471 U.S. Marines: A genome-wide association study across multiple ancestries implicates
472 PRTFDC1 as a potential PTSD gene. *Psychoneuroendocrinology* **51**, 459–471 (2015).
- 473 31. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of
474 genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
- 475 32. Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of
476 complex trait heritability. *Nat Genet* **52**, 1355–1363 (2020).
- 477 33. Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-
478 wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).
- 479 34. Powers, A. *et al.* The differential effects of PTSD, MDD, and dissociation on CRP in
480 trauma-exposed women. *Comprehensive Psychiatry* **93**, 33–40 (2019).
- 481 35. Beck, J. G., Grant, D. M., Clapp, J. D. & Palyo, S. A. Understanding the interpersonal
482 impact of trauma: Contributions of PTSD and depression. *Journal of Anxiety Disorders* **23**,
483 443–450 (2009).
- 484 36. Adams, S. W. *et al.* PTSD and Comorbid Depression: Social Support and Self-Efficacy in
485 World Trade Center Tower Survivors 14–15 Years After 9/11. *Psychol Trauma* **11**, 156–164
486 (2019).

- 487 37. Hovens, J. G. F. M. *et al.* Childhood life events and childhood trauma in adult patients with
488 depressive, anxiety and comorbid disorders vs. controls. *Acta Psychiatrica Scandinavica*
489 **122**, 66–74 (2010).
- 490 38. Lewis, S. J. *et al.* The epidemiology of trauma and post-traumatic stress disorder in a
491 representative cohort of young people in England and Wales. *The Lancet Psychiatry* **6**, 247–
492 256 (2019).
- 493 39. Bremner, J. D., Southwick, S. M., Darnell, A. & Charney, D. S. Chronic PTSD in Vietnam
494 combat veterans: Course of illness and substance abuse. *The American Journal of Psychiatry*
495 **153**, 369–375 (1996).
- 496 40. Korem, N., Duek, O., Xu, K., Harpaz-Rotem, I. & Pietrzak, R. H. Cannabinoid Receptor 1
497 rs1049353 Variant, Childhood Abuse, and the Heterogeneity of PTSD Symptoms: Results
498 From the National Health and Resilience in Veterans Study. *Chronic Stress* **5**,
499 24705470211011076 (2021).
- 500 41. Brittain, J. M. *et al.* Suppression of inflammatory and neuropathic pain by uncoupling
501 CRMP-2 from the presynaptic Ca²⁺ channel complex. *Nat Med* **17**, 822–829 (2011).
- 502 42. Fukata, Y. *et al.* CRMP-2 binds to tubulin heterodimers to promote microtubule assembly.
503 *Nat Cell Biol* **4**, 583–591 (2002).
- 504 43. Khanna, R. *et al.* Opening Pandora’s jar: a primer on the putative roles of CRMP2 in a
505 panoply of neurodegenerative, sensory and motor neuron, and central disorders. *Future*
506 *Neurology* **7**, 749–771 (2012).
- 507 44. Kékesi, K. A. *et al.* Altered Functional Protein Networks in the Prefrontal Cortex and
508 Amygdala of Victims of Suicide. *PLOS ONE* **7**, e50532 (2012).

- 509 45. Natalia Silva, P. *et al.* CNP and DPYSL2 mRNA Expression and Promoter Methylation
510 Levels in Brain of Alzheimer's Disease Patients. *Journal of Alzheimer's Disease* **33**, 349–
511 355 (2013).
- 512 46. Williamson, R. *et al.* CRMP2 Hyperphosphorylation is Characteristic of Alzheimer's
513 Disease and not a Feature Common to Other Neurodegenerative Diseases. *Journal of*
514 *Alzheimer's Disease* **27**, 615–625 (2011).
- 515 47. Lee, H. *et al.* Changes in Dpysl2 expression are associated with prenatally stressed rat
516 offspring and susceptibility to schizophrenia in humans. *International Journal of Molecular*
517 *Medicine* **35**, 1574–1586 (2015).
- 518 48. Pham, X. *et al.* The DPYSL2 gene connects mTOR and schizophrenia. *Transl Psychiatry* **6**,
519 e933–e933 (2016).
- 520 49. Zhao, X. *et al.* An investigation of the dihydropyrimidinase-like 2 (DPYSL2) gene in
521 schizophrenia: genetic association study and expression analysis. *International Journal of*
522 *Neuropsychopharmacology* **9**, 705–712 (2006).
- 523 50. Izumi, R. *et al.* Dysregulation of DPYSL2 expression by mTOR signaling in schizophrenia:
524 Multi-level study of postmortem brain. *Neuroscience Research* (2021)
525 doi:10.1016/j.neures.2021.09.004.
- 526 51. Fallin, M. D. *et al.* Linkage and association on 8p21.2-p21.1 in schizophrenia. *American*
527 *Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **156**, 188–197 (2011).
- 528 52. Nakata, K. *et al.* The human dihydropyrimidinase-related protein 2 gene on chromosome
529 8p21 is associated with paranoid-type schizophrenia. *Biological Psychiatry* **53**, 571–576
530 (2003).

- 531 53. Taylor, A. & Wang, K.-S. Association between DPYSL2 gene polymorphisms and alcohol
532 dependence in Caucasian samples. *J Neural Transm* **121**, 105–111 (2014).
- 533 54. Lai, C. *et al.* CD8+CD103+ tissue-resident memory T cells convey reduced protective
534 immunity in cutaneous squamous cell carcinoma. *J Immunother Cancer* **9**, e001807 (2021).
- 535 55. Banchereau, R. *et al.* Intratumoral CD103+ CD8+ T cells predict response to PD-L1
536 blockade. *J Immunother Cancer* **9**, e002231 (2021).
- 537 56. Hu, X. *et al.* ITGAE Defines CD8+ Tumor-Infiltrating Lymphocytes Predicting a better
538 Prognostic Survival in Colorectal Cancer. *EBioMedicine* **35**, 178–188 (2018).
- 539 57. Rodrigo, L. M. & Nyholt, D. R. Imputation and Reanalysis of ExomeChip Data Identifies
540 Novel, Conditional and Joint Genetic Effects on Parkinson’s Disease Risk. *Genes* **12**, 689
541 (2021).
- 542 58. Lesch, K.-P. *et al.* Molecular genetics of adult ADHD: converging evidence from genome-
543 wide association and extended pedigree linkage studies. *J Neural Transm* **115**, 1573–1585
544 (2008).
- 545 59. Zhou, K. *et al.* Meta-analysis of genome-wide linkage scans of attention deficit hyperactivity
546 disorder. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1392–1398 (2008).
- 547 60. Maj, C. *et al.* Intermediate lengths of the C9ORF72 hexanucleotide repeat expansion may
548 synergistically contribute to attention deficit hyperactivity disorder in child and his father:
549 case report. *Neurocase* **27**, 138–146 (2021).
- 550 61. Chivers-Wilson, K. A. Sexual assault and posttraumatic stress disorder: A review of the
551 biological, psychological and sociological factors and treatments. *McGill J Med* **9**, 111–118
552 (2006).

- 553 62. Dunmore, E., Clark, D. M. & Ehlers, A. A prospective investigation of the role of cognitive
554 factors in persistent Posttraumatic Stress Disorder (PTSD) after physical or sexual assault.
555 *Behaviour Research and Therapy* **39**, 1063–1084 (2001).
- 556 63. Ullman, S. E. & Filipas, H. H. Predictors of PTSD Symptom Severity and Social Reactions
557 in Sexual Assault Victims. *J Trauma Stress* **14**, 369–389 (2001).
- 558 64. Steenkamp, M. M., Dickstein, B. D., Salters-Pedneault, K., Hofmann, S. G. & Litz, B. T.
559 Trajectories of PTSD symptoms following sexual assault: Is resilience the modal outcome?
560 *Journal of Traumatic Stress* **25**, 469–474 (2012).
- 561 65. Johnson, D. M. & Zlotnick, C. HOPE for battered women with PTSD in domestic violence
562 shelters. *Prof Psychol Res Pr* **40**, 234–241 (2009).
- 563 66. Maihofer, A. X. *et al.* Enhancing Discovery of Genetic Variants for Posttraumatic Stress
564 Disorder Through Integration of Quantitative Phenotypes and Trauma Exposure Information.
565 *Biological Psychiatry* **0**, (2021).
- 566 67. Sartor, C. E. *et al.* Common Heritable Contributions to Low-Risk Trauma, High-Risk
567 Trauma, Posttraumatic Stress Disorder, and Major Depression. *Archives of General*
568 *Psychiatry* **69**, 293–299 (2012).
- 569 68. Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A. & Livesley, W. J. Genetic and
570 Environmental Influences on Trauma Exposure and Posttraumatic Stress Disorder
571 Symptoms: A Twin Study. *AJP* **159**, 1675–1681 (2002).
- 572 69. Nievergelt, C. M. *et al.* Genomic Approaches to Posttraumatic Stress Disorder: the
573 Psychiatric Genomic Consortium Initiative. *Biol Psychiatry* **83**, 831–839 (2018).
- 574
575

576 **Figure Legends**

577

578 **Figure 1. Trauma risk score (TRS) category interactions and averages among PTSD cases**
579 **and controls.** (A) Interactions between Traumatic/Stressful Life Events (TSLEs) of TRS
580 categories. (B) Average TRS (weighted by elastic net regression beta values) between PTSD
581 cases and controls and (C) average TRS (weighted by frequency of trauma exposure) between
582 PTSD cases and controls.

583

584

585

