nature research

Corresponding author(s): Joel Gelernter, Murray B. Stein

Last updated by author(s): May 27, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Details on PTSD phenotyping in the MVP cohort have been described previously, see Harrington et al. 2019 (PMID: 31009556). For the re-experiencing phenotype specifically, only SAS software was used
Data analysis	Minimac3, PLINK, flashpca, RVTEST, MAGMA, FUMA, LD score regression, LDHub v2.0, PRSice, GTEx Analysis Release V7 - eQTL, UK Biobank - summary statistics, 1,000 Genomes Project Reference data, R packages (GenABEL, gtx, psych), STRING, and Ingenuity IPA (Ingenuity Systems, Redwood City, CA, USA).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Summary statistics will be available via dbGaP. The dbGaP accession assigned to the Million Veteran Program is phs001672.v1.p1. The website is: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001672.v1.p1

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined, but instead reflects our best effort to aggregate all possible invidividuals with genome-wide genotype data and robust phenotyping of re-experiencing symptoms included in the PTSD checklist for DSM-4 (PCL-4) available in the MVP dataset at the time of the analysis (August 2018).
Data exclusions	As of August 2018, more than 670,000 veterans have enrolled in the program; for the current analyses, genotyping data were available from approximately 350,000 participants. After accounting for missing phenotype data as only subjects who completed the voluntary surveys could be included and excluding individuals with poor genetic information, the final sample sizes were 146,660 participants of European descent and 19,983 participants of African descent.
Replication	We used UKB summary association data (117,900 subjects of European descent) for an item from the traumatic events assessment which is one of the five REX items in MVP: "Repeated disturbing thoughts of stressful experience in past month" (UKB Field ID: 20497). We tested different levels of replication. We observed five single-variant replications out of eight GWS loci identified in the MVP cohort. We also observed a significant genetic correlation between MVP and UKB datasets. The genetic overlap was also confirmed using a polygenic risk score approach.
Randomization	Randomization of experimental groups was not applicable to this study. The experimental conditions are determined by each individual's genetics, which are fixed at conception. Conceptually this reflects a randomization of the alleles inherited from each individual's parents (i.e. mendelian randomization), but it does not involve randomization of experimental conditions by the researchers in a classical sense. Our study assess the observed association between that natural randomization of genotype and the ascertained phenotype of re-experiencing symptoms.
Blinding	Blinding is not relevant to the current study. Samples were not allocated to different conditions by the researchers, and the phenotype ascertainment process is fully separate from the genotyping process.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The demographic characteristics are reported in Table S2 and they include age, sex, ethnicity, marital status, ,military branch, and service era. Genetic information was used to confirm the ancestry of the participants using 1000 Genomes populations as reference populations. We considered the genetic principal components derived from the genetic information together with age and sex as covariates in the genetic association analysis, which was stratified by ancestry group.
Recruitment	All subjects are enrollees in the MVP. Active users of the Veterans Health Administration healthcare system (>8 million veterans) learn of MVP via an invitational mailing and/or through MVP staff while receiving clinical care with informed consent and HIPAA authorization as the only inclusion criteria. Enrollment involves providing a blood sample for genomic analyses, allowing ongoing access to medical records and other administrative health data by authorized MVP staff, and completing questionnaires. Two surveys are provided to MVP participants: the MVP Baseline Survey, for demographic factors, family pedigree, health status,

	lifestyle habits, military experiences, medical history, family history of specific illnesses, and physical features; and the optional MVP Lifestyle Survey, which includes the PCL (DSM-IV version).
Ethics oversight	Research involving MVP in general is approved by the VA Central IRB; the current project was also approved by IRBs in Boston, San Diego, and West Haven.

Note that full information on the approval of the study protocol must also be provided in the manuscript.