***Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder***

**Supplementary Material**

**Methods**

**Study sample and phenotype definitions**

***UK Biobank***

Between 2006 and 2010, the UK Biobank recruited over 500,000 individuals for a cohort study aimed at improving diagnosis, treatment and prevention of serious diseases[(Allen *et al.*, 2014)](https://paperpile.com/c/d3jKGH/DU6Pp). Participants gave full informed consent, answered surveys and provided physical measurements including DNA samples at a baseline visit to one of 22 assessment centres across the UK [(Sudlow *et al.*, 2015)](https://paperpile.com/c/d3jKGH/8B7V9). The phenotypes assessed in this study were derived from the online follow-up Mental Health Questionnaire (MHQ), which received 157,366 responses. This online questionnaire comprises a number of adapted versions of clinically used psychiatric questionnaires to assess common mental health disorders. Case definitions based on responses to the psychiatric questionnaires were derived by the working committee who wrote the MHQ [(Davis *et al.*, 2020)](https://paperpile.com/c/d3jKGH/StIv). The individual-level analysis sample for the present study specifically focused on participants who had completed the MHQ and met criteria for lifetime major depressive disorder (MDD) (*N*=29,41).

*Major depressive disorder (MDD)*

Participants were considered cases for probable MDD based on their responses to questions derived from the Composite International Diagnostic Interview Short Form (CIDI-SF). Reporting on a period of depression lasting at least two weeks, cases endorsed at least one of the two core symptoms (“ever had prolonged feelings of sadness or depression” and “ever had prolonged loss of interest in normal activities”), at least five of the nine symptoms queried overall, and reported that they were affected almost every day, most days during the period, with more than a little impact on normal functioning. Controls did not meet case criteria and did not meet criteria for a current episode of depression. Participants who self-reported a diagnosis of schizophrenia, other psychoses, or bipolar disorder were excluded. Controls were excluded if they self-reported any mental illness, taking any drug with an antidepressant indication, or had been hospitalised with a mood disorder or met previously defined criteria for a mood disorder.

*Major depressive disorder with/without reported psychological trauma*

Participants were classified as having either “MDD with reported exposure to psychological trauma” or “MDD without reported exposure psychological trauma”. Phenotype definitions for these are included in the Supplementary Material of Coleman et al. (2020). Questions relating to traumatic experiences in childhood were assessed on a five point scale (ranging from “never” to “often”) using the Childhood Trauma Screener (an adapted version of the Childhood Trauma Questionnaire [(Bernstein *et al.*, 1994; Grabe *et al.*, 2012; Bellis *et al.*, 2014)](https://paperpile.com/c/d3jKGH/tGYdL%2BRodou%2ByggKH). An equivalent screener was constructed for traumatic events in adulthood. Only traumatic experiences with an odds ratio >2.5 with MDD were selected to obtain a single binary variable for trauma exposure. This included: three events in childhood (did not feel loved, felt hated by a family member, sexually abused); three events in adulthood (physical violence, belittlement, sexual interference); and one PTSD-related event (ever a victim of sexual assault). Participants were included in the “MDD with reported exposure to psychological trauma” analysis if they reported two or more of these events and met criteria for MDD. Participants were included in the “MDD without reported exposure to psychological trauma” analysis if they reported none of these events and met case criteria for MDD [(Coleman *et al.*, 2020)](https://paperpile.com/c/d3jKGH/ZUvoZ).

*Recurrent and single-episode major depressive disorder*

Participants were classified as having either recurrent MDD or single-episode MDD. The definitions of the recurrent and single-episode MDD phenotypes can be found in the Supplementary Material of Coleman et al. (2019). In brief, participants who met criteria for MDD were classified with recurrent MDD if they reported multiple depressed periods across their lifetime and single-episode MDD otherwise [(Coleman *et al.*, 2019)](https://paperpile.com/c/d3jKGH/8iWU4).

***Posttraumatic stress disorder phenotypes***

*Posttraumatic stress disorder in the UK Biobank*

The definition of posttraumatic stress disorder (PTSD) in the UK Biobank is included in the Supplementary Material of Nievergelt et al. (2019). In brief, the PTSD phenotype was derived from six questions asked in the follow-up online mental health questionnaire. These questions were derived from the brief civilian version of the PTSD Checklist Screener (PCL-S) which measures PTSD symptoms experienced in the previous month: avoidance of activities; disturbing thoughts; and feeling upset; and two additional questions related to feeling distant and feeling irritable. Each item was scored on a five-point Likert item measuring the amount of concern caused by that symptom in the past month (1="Not at all" to 5="Extremely"). In addition, a “trouble concentrating” question from the Patient Health Questionnaire-9 (PHQ9) depression questionnaire was added to replace a similar item that would normally be included in the PCL-S. This item was scored on a four-point Likert item according to frequency of difficulties associated with trouble concentrating (1="Not at all" and 4="Nearly every day"). For each participant, all items were summed into a total score ranging 3-29. Participants were considered PTSD cases if they had an overall PCL-S score ≥ 13. Participants were considered PTSD controls if they responded to all of the initial three questions and had PCL-S score ≤ 12 [(Nievergelt *et al.*, 2019)](https://paperpile.com/c/d3jKGH/Va9XV).

 *Posttraumatic Stress Disorder working group of the Psychiatric Genomics Consortium*

The PTSD working group of the Psychiatric Genomics Consortium (PGC) meta-analysed data from 59 studies of PTSD to perform a GWAS known as the PTSD Freeze 1.5 (PGC1.5-PTSD). This sample involved 12,823 cases and 35,648 controls. The PGC gathered data for PGC1.5-PTSD through a number of independent studies who used a wide range of methods, primarily telephone diagnostic interviews and face-to-face clinical assessments. Some of the participants included in this cohort are veterans who have been combat or war-zone exposed. Other traumatic events assessed by the PGC include serious car accidents, campus shootings, domestic violence, and childhood physical and sexual abuse. Participants were assessed for current and lifetime PTSD using various instruments and different versions of the DSM [(Nievergelt *et al.*, 2019)](https://paperpile.com/c/d3jKGH/Va9XV). Further details of the contributing studies and instruments used to assess PTSD are contained in the Supplementary Material of Nievergelt et al. (2019).

To maximise power in genome-wide analysis, the UK Biobank and PTSD Freeze 1.5 (PGC1.5-PTSD) were combined. This combined data set is known as the PTSD Freeze 2 (PGC2-PTSD). In this sample, the number of cases was 23,212 and the number of controls was 151,447. The SNP-based heritability (liability scale) of the PGC2-PTSD phenotype was 0.06 (SE=0.011). In the main paper, we present genetic correlations between the MDD categories and the PGC1.5-PTSD summary statistics. We also calculated genetic correlations with the PGC2-PTSD summary statistics which are presented in Supplementary Table 1.

*The Million Veteran Program*

The Million Veteran Program (MVP) is a longitudinal study of United States military veterans who provided a blood sample for biobanking and responses to various questionnaires [(Gaziano *et al.*, 2016)](https://paperpile.com/c/d3jKGH/5qmKA). While the MVP sample consists of United States veterans, only 27.5% have confirmed war- or combat- exposure, while 29.3% who had not been exposed. The remaining 43.1% had unknown war- and combat-exposure [(Stein *et al.*, 2020)](https://paperpile.com/c/d3jKGH/CWNcF). The MVP-PTSD GWAS was of a binary, algorithmically-defined probable PTSD phenotype based on the veterans’ Electronic Health Records. Further details of how the MVP phenotype was defined can be found in Stein et al. (2020).

 **Computational Methods**

***High Definition Likelihood inference of genetic correlations***

Genetic correlations were estimated using High Definition Likelihood (HDL). Firstly for each phenotype, GWAS summary statistics were used to estimate the proportion of variance explained by common genetic variants (h2SNP) using High Definition Likelihood (HDL). While this is not a specific aim of the study, this step is necessary for interpreting genetic correlations between the PTSD phenotypes and the four MDD categories. HDL rests upon the principle of linkage-disequilibrium and extends the regression formula used by Linkage Disequilibrium Score Regression (LDSC) (see below). Unlike LDSC, HDL uses a full, likelihood-based method to estimate genetic correlations [(Ning *et al.*, 2020)](https://paperpile.com/c/d3jKGH/rx2BZ). Further details can be found in the original paper by Ning et al. (2020).

All summary statistics were wrangled using the built-in HDL function for data wrangling (<https://github.com/zhenin/HDL/wiki/Format-of-summary-statistics>). All GWAS summary statistics used in this analysis had at least 99% SNP overlap with the UK Biobank LD reference panel using HapMap3 variants, apart from the Million Veteran Program (MVP) GWAS which had an overlap of 94.23%. Through correspondence with the HDL authors, we are confident that this level of SNP overlap is acceptable since the mismatch is not due to differences in the ancestral population from which the samples were created (e.g European vs. non-European populations) since both the MVP and the UK Biobank samples comprise participants from European ancestries only. The authors of HDL have performed a simulation to test this and discovered that missing SNPs lead to more conservative results but should not generate false positives (Z Ning, personal correspondence). Therefore, a missing rate around 5% is acceptable. Furthermore, we repeated the data wrangling using the smaller HapMap2 (as opposed to HapMap3) reference panel and this did not improve the SNP overlap appreciably (data not shown). The authors of HDL therefore advised us that we continue to use the HapMap3 reference panel for the wrangling of the MVP Case-Control GWAS summary statistics.

***High Definition Likelihood block-jackknife***

The block-jackknife method was used to compare genetic correlations for a statistically significant difference between them. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations within each set (the sets being UKB-PTSD, PGC2-PTSD, and MVP-PTSD presented in the main paper and PGC1.5-PTSD presented in the Supplementary Material). The block-jackknife uses a resampling method to estimate standard errors for each genetic correlation, which is then used to determine whether the differences between the genetic correlations are significantly different to zero. Here we provide an explanation of the block-jackknife method (Z Ning, personal correspondence).

When estimating genetic correlations, *r*g1 and *r*g2, one may then want to test whether the difference between them, referred to as the “global difference”, is significantly different to zero. This would mean the null hypothesis is: *r*g1 - *r*g2 = 0. For *r*g1, by setting jackknife.df=TRUE when using the HDL tool, you can get jackknife estimates of *r*g1.1 to *r*g1.61. This creates a file with 61 jackknife estimates because the genome is split into 61 pieces during the resampling process. *r*g1.k represents the estimated *r*g with piece k removed. Similarly, you have *r*g2.1 to *r*g2.61 for *r*g2.

With these two files containing the values in for both correlations, one can create the block-jackknife estimates by doing *r*g1.1 - *r*g2.1 all the way up to *r*g1.61 - *r*g2.61. The block-jackknife standard error can be found using this formula:

Where = 61 (for each of the resampling estimates), = the difference between each resampling estimate (e.g., *r*g1.1 - *r*g2.1 all the way up to *r*g1.61 - *r*g2.61), = the mean of these estimates.

Then, the block-jackknife standard error and the global difference between the two correlations can be used to perform a Wald test:

Z = global difference block-jackknife standard error

The Wald test will give a z-score, which can then be used to find the two-tailed p-value, which will tell you whether the global difference is significantly different to zero. In our study, we corrected for multiple testing by considering *p*<0.008 as the threshold for significance (0.05/6 = 0.008, to account for the 6 block-jackknife tests carried out for each PTSD phenotype). Supplementary Table 5 contains all results from the HDL block-jackknife analysis.

***Converting observed scale heritability estimates to the liability scale***

The heritability of each trait was estimated by the HDL programme irrespective of population prevalence. Therefore, these estimates were converted to the liability scale in R, using code from the Nievergelt Lab github: https://gist.github.com/nievergeltlab/fb8a20feded72030907a9b4e81d1c6ea. Standard errors were also converted to the liability scale using the same formula[(Lee *et al.*, 2011)](https://paperpile.com/c/d3jKGH/nJmyE). Population prevalence for PTSD was adopted from Nievergelt et al. (2019), from Coleman et al. (2020) for reported trauma in MDD and from Burcusa & Iacono (2007) for recurrence in MDD.

***Linkage Disequilibrium Score Regression***

In addition to using HDL, we calculated genetic correlations using Linkage Disequilibrium Score Regression (LDSC). LDSC is a command line tool for estimating heritability and genetic correlation from GWAS summary statistics. It rests upon the principle of linkage-disequilibrium (LD). LD describes the degree to which an allele of one SNP is inherited or correlated with an allele of another SNP within a population [(Bush and Moore, 2012)](https://paperpile.com/c/d3jKGH/7ET8w). In this method, an “LD Score” of a given SNP refers to “the sum of LD r2 measured with all other SNPs”. LDSC works by performing regression analysis on the LD scores and the test statistic of each SNP included in the GWAS, including those that do not meet genome-wide significance [(Bulik-Sullivan *et al.*, 2015a)](https://paperpile.com/c/d3jKGH/pjrcr). LDSC relies on the fact that the GWAS effect-size estimate for a given SNP incorporates the effect of all SNPs in LD with that particular SNP. For most complex human traits, which are polygenic, SNPs with high LD will have higher chi-square test statistics, on average, than SNPs with low LD [(Bulik-Sullivan *et al.*, 2015b)](https://paperpile.com/c/d3jKGH/Kylad).

As in the HDL analysis, GWAS summary statistics were used to estimate the proportion of variance explained by common genetic variants (h2SNP) using LDSC. While this is not a specific aim of the study, this step is necessary for interpreting genetic correlations.

***Linkage Disequilibrium Score Regression (LDSC) block-jackknife***

The block-jackknife method was used to compare genetic correlations for a statistically significant difference between them. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations within each set (the sets being UKB-PTSD, PGC1.5-PTSD, PGC2-PTSD, and MVP-PTSD). As described above for HDL, the block-jacknife works by repeated re-estimation of blocks of jack-knife estimates. In the case of LDSC, the number of blocks is set by the user (we used 200), but otherwise the calculation of significant differences follows the description provided for HDL.

***Polygenic Risk Scores***

PRS were calculated PRSice v2.3.1, a command line programme that uses GWAS summary statistics to calculate genetic risk of a base phenotype in individuals from an independent sample. A PRS refers to the summation of alleles across many genetic loci associated with a particular trait or disease. These alleles are typically weighted by effect sizes estimated from GWAS [(Euesden *et al.*, 2015)](https://paperpile.com/c/d3jKGH/PRbBX). In our study, the PRS represents the aggregated PTSD risk conferred by many DNA variants in participants of the UK Biobank who have MDD. Since even well-powered GWAS offer only tentative evidence for causally associated variants, PRS are calculated at a range of different *P*-value thresholds to provide the ‘best-fit’, or most predictive, PRS [(Dudbridge, 2013)](https://paperpile.com/c/d3jKGH/8DMtH). Once the best fitting PRS has been estimated, these are used as predictors of a target phenotype in individuals in an independent sample in a regression. PRS were calculated at 11 p-value thresholds (5x10-8, 1x10-5, 1x10-3, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1). Phenotype permutations (implemented in PRSice) were used to produce an empirical p-value for the association at the best-fitting PRS, which accounts for testing multiple thresholds. We then performed a logistic regression to examine whether the risk scores are more strongly associated with MDD with reported trauma or MDD without reported trauma, and recurrent or single-episode MDD.

***Power calculations***

We calculated the power of our PRS analyses using the Additive Variance Explained and Number of Genetic Effects Method of Estimation (AVENGEME) programme in R [(Dudbridge, 2013)](https://paperpile.com/c/d3jKGH/8DMtH). We were uncertain about the covariance between genetic effects in the target sample (MVP PTSD) and the two target samples. Therefore, based on the parameters required by the AVENGEME package (see below), we calculated that for the PRS analysis testing the risk score’s association with MDD with and without reported trauma to be powered at least 80%, the covariance between the genetic effect sizes in the training and target samples would need to be at least 0.024. We calculated that for the PRS analysis testing the risk score’s association with recurrent and single-episode MDD to be powered at least 80%, the covariance between the genetic effect sizes in the training and target samples would need to be at least 0.0305. Due to these minimum covariance estimates being so low, we were confident that the PRS analyses were powered by at least 80%.

*Parameters required by the AVENGEME package:*

● Number of SNPs included in PRS after clumping = 183,881

● Proportion of variance in PTSD explained by genetic effects = 0.03 [(Stein *et al.*, 2020)](https://paperpile.com/c/d3jKGH/CWNcF)

● Training sample prevalence of PTSD = 0.18 [(Stein *et al.*, 2020)](https://paperpile.com/c/d3jKGH/CWNcF)

● Population prevalence of PTSD = 0.30 [(Nievergelt *et al.*, 2019)](https://paperpile.com/c/d3jKGH/Va9XV)

● Target sample prevalence of reported trauma among MDD cases = 0.59

● Population prevalence of reported trauma among MDD cases = 0.52 [(Coleman *et al.*, 2020)](https://paperpile.com/c/d3jKGH/ZUvoZ)

● Target sample prevalence of recurrent MDD = 0.59

● Population prevalence of recurrent MDD among MDD cases = 0.5 [(Burcusa and Iacono, 2007)](https://paperpile.com/c/d3jKGH/frZy3)

**Results**

**High Definition Likelihood block-jackknife**

The results of the HDL block-jackknife analysis can be found in Supplementary Table 5. Differences between genetic correlations were considered statistically significant if they surpassed the Bonferroni corrected alpha (0.05/6 = 0.008; i.e. to correct for the six block-jackknife tests per PTSD phenotype).

**Linkage Disequilibrium Score Regression genetic correlations**

Genetic correlations (Supplementary Table 6) estimated by LDSC were considered significantly different to zero and to one if they reached or surpassed the Bonferroni-corrected alpha in each analysis (0.05/4 = 0.0125; i.e. to correct for the four tests).

**Linkage Disequilibrium Score Regression block-jackknife**

We tested the differences between the genetic correlations using a block-jackknife. Differences were considered statistically significant if they passed a Bonferroni-corrected alpha of 0.008 (0.05/6 = 0.008; i.e. to correct for the six block-jackknife tests per PTSD phenotype). Supplementary Table 7 contains the results of the LDSC block-jackknife.

***Comparison of the two methods for estimating genetic correlations***

In our study, we performed the genetic correlation and block-jackknife analyses in both HDL and LDSC. We have opted to only present the HDL results in the main text since it is the preferable method (for reasons discussed in the paper). Nonetheless, there are some key differences between the HDL and LDSC results which need to be discussed here.

An anticipated difference was that no genetic correlations were found to differ significantly from any other genetic correlations when using LDSC. The absence of statistically significant differences in genetic correlation must be interpreted in the context of the power of the original GWAS from which the summary statistics used in this study were created. As shown in Supplementary Table 6, the standard errors surrounding the LDSC point estimates are notably large, meaning that any differences between genetic correlation would need to be large to be detected as significantly different to zero in the block-jackknife analysis. The advantage of using HDL to estimate genetic correlations is reduction in variance of the point estimate [(Ning *et al.*, 2020)](https://paperpile.com/c/d3jKGH/rx2BZ), which provides better power to observe differences between genetic correlations. This was the case in our study, where we found that PTSD is significantly more genetically correlated with recurrent MDD than it is genetically correlated with MDD without reported trauma when using HDL. Similar to the HDL results, when using LDSC we find that all PTSD phenotypes have a higher genetic correlation with recurrent MDD than with MDD without reported trauma; however, unlike when using the HDL block-jackknife, this difference was not significantly different to zero in the results of LDSC block-jackknife.

Another difference between the HDL and LDSC results is the pattern of genetic correlations. Firstly, when using LDSC, we observe a clear pattern where all PTSD phenotypes are more genetically correlated with recurrent MDD than single-episode MDD. We do not see this pattern when using HDL. Secondly, the consistent pattern whereby all PTSD phenotypes are more genetically correlated with MDD with reported trauma compared MDD without reported trauma when using HDL only holds true for the PGC 1.5, PGC 2 and MVP PTSD phenotypes when using LDSC, whereas the UK Biobank PTSD phenotype appears to be slightly more genetically correlated with MDD with without reported trauma (*r*g difference = 0.0039).

These differences are likely due to inconsistent SNP reference panels used by HDL and LDSC, which are needed to estimate the LD-scores. Therefore, a difference in the reference panel between the two methods is likely to lead to differing results. In HDL, the reference panels with imputed SNPs are based on genotypes in UK Biobank, which were imputed to HRC and UK10K + 1000 Genomes. Specifically in our study, we used the 1,029,876 Quality Controlled UK Biobank imputed HapMap3 SNPs reference panel (<https://github.com/zhenin/HDL/wiki/Reference-panels>). On the other hand, LDSC uses a reference panel from the 1000 Genomes Project which is not specific to the UK Biobank [(Bulik-Sullivan *et al.*, 2015b)](https://paperpile.com/c/d3jKGH/Kylad). According to correspondence with the authors, HDL uses the UK Biobank reference panels instead of 1000 Genomes because the UK Biobank has a larger sample size. This therefore leads to more accurate estimates of LD. Further details can be found at the discussion section of Ning et al. (2020). However, this increase in the accuracy of LD estimation is likely to apply most strongly to summary statistics from the UK in general and UK Biobank in particular, and less so to samples descended from European ancestry populations from elsewhere in Europe. This being the case, it is feasible that the differences in genetic correlations we observe in our study partly reflect differences in the proportion of UK ancestry in the PTSD summary statistics (as all of the MDD summary statistics were drawn from UK Biobank). This inconsistency between LDSC and HDL, and its potential relationship to UK ancestry in the summary statistics assessed, is likely to have wider implications than our study alone and requires a detailed examination beyond the scope of this study.

**Polygenic risk scores**

Polygenic risk scores for PTSD were calculated PRSice v2.3.1. Two PRS analyses were performed: two regressions using PRS based on the MVP PTSD summary statistics. Full results of these analyses can be found in Supplementary Table 8. Beta coefficients were exponentiated in R to give odds ratios (OR) and 95% confidence intervals (CI).

***Nagelkerke’s R2***

The PRS were initially calculated without consideration of the prevalence of the target phenotype (reported psychological trauma in individuals with MDD and episode recurrence in individuals with MDD) within the population. Based on the results from PRSice, Nagelkerke’s R2 was calculated for the estimated population prevalence, +10% and -10% in R. We obtained an estimated population prevalence for trauma exposure among MDD cases of 52% from the Supplementary Material of Coleman et al. (2020). We obtained an estimated population prevalence for recurrence among MDD cases of 50% from Burcusa & Iacono (2007). For the regression testing the PTSD risk scores’ association with MDD with reported trauma and MDD without reported trauma, the Nagelkerke’s R2 based on a population prevalence of 42%, 52% and 62% were 0.0563%, 0.0569% and 0.0555% respectively. For the regression testing the PTSD risk scores’ association with recurrent MDD compared to single-episode MDD, the Nagelkerke’s R2 based on population prevalence 40%, 50% and 60% were 0.0296%, 0.0301% and 0.0296% respectively.

**Supplementary Tables**

Supplementary Table 1: High Definition Likelihood (HDL) genetic correlation estimates (*r*g), standard errors (SE) and 95% confidence intervals (CI) of Psychiatric Genomics Consortium 2 PTSD (PGC2-PTSD) with the four major depressive disorder (MDD) categories. *P* (diff 0) refers to *p*-value to test whether the *r*g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the *r*g estimate differs from 1. Genetic correlations were considered significant if they reached or surpassed the Bonferroni adjusted threshold (*p*<0.0125). Significant *p*-values are shown in bold.The SNP-based heritability of the PGC2-PTSD phenotype was estimated by HDL to be 0.06 (SE = 0.006).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| PTSD Phenotype | MDD phenotype | *r*g | SE | Lower CI | Upper CI | *P* (diff 0) | *P* (diff 1) |
| PGC2-PTSD | MDD with reported trauma | 0.6497 | 0.0825 | 0.4880 | 0.8114 | **3.39x10-15** | **2.18x10-5** |
| PGC2-PTSD | MDD without reported trauma | 0.5509 | 0.1107 | 0.3339 | 0.7679 | **5.12x10-7** | **4.24x10-5** |
| PGC2-PTSD | Recurrent MDD | 0.7915 | 0.0821 | 0.6306 | 0.9524 | **8.45x10-22** | **0.01** |
| PGC2-PTSD | Single-episode MDD | 0.8147 | 0.1129 | 0.5934 | 1.0360 | **5.27x10-13** | 0.1 |

Supplementary Table 2: Difference in reporting rates of traumatic life events, assessed by the Mental Health Questionnaire (MHQ), between individuals with recurrent and single-episode major depressive disorder (MDD) in UK Biobank MHQ respondents who met criteria for lifetime MDD (*N*=29,471). Traumatic events include two childhood events, two adulthood events and five catastrophic/posttraumatic stress disorder (PTSD)-related events. Differences were considered significant if they surpassed the Bonferroni adjusted alpha (*p*<0.006) to correct for the nine chi-square tests. Significant differences are shown in bold.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Trauma category | Traumatic event | Endorsement in single-episode MDD (%) | Endorsement in recurrent MDD (%) | 2 statistic | *P*-value |
| Childhood physical abuse | Physically abused by family as a child | 2,644 (22%) | 4,846 (28%) | 126 | **3.82x10-29** |
| Childhood physical neglect | Someone to take to doctor when needed as a child | 1,948 (16%) | 3,556 (20%) | 80 | **3.80x10-19** |
| Adulthood emotional neglect | Been in a confiding relationship as an adult | 3,758 (32%) | 6,937 (40%) | 205 | **2.08x10-46** |
| Adulthood physical neglect | Able to pay rent/mortgage as an adult | 1,963 (17%) | 3,603 (21%) | 87 | **1.07x10-20** |
| PTSD-related: experience of war or combat | Been involved in combat or exposed to war-zone | 400 (3%) | 595 (3%) | 0.13 | 0.72 |
| PTSD-related: serious accident | Been in serious accident believed to be life-threatening | 1,247 (10%) | 2,313 (13%) | 55 | **9.47x10-14** |
| PTSD-related: life-threatening illness | Diagnosed with life-threatening illness | 2,092 (17%) | 3,211 (18%) | 5 | 0.03 |
| PTSD-related: physically violent crime | Victim of physically violent crime | 2,332 (19%) | 4,273 (25%) | 106 | **8.12x10-25** |
| PTSD-related: witnessed sudden violent death | Witnessed sudden violent death | 1,778 (15%) | 2,866 (16%) | 14 | **0.0002** |

Supplementary Table 3: High Definition Likelihood (HDL) genetic correlation estimates (*r*g), standard errors (SE), and 95% confidence intervals (CI) between the four posttraumatic stress disorder (PTSD) phenotypes: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). *P* (diff 0) refers to *p*-value to test whether the *r*g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the *r*g estimate differs from 1. Genetic correlations were considered significant if they surpassed the Bonferroni adjusted threshold (*p*<0.008) to correct for the 6 tests. Significant *p*-values are shown in bold.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Phenotype 1** | **Phenotype 2** | ***r*g** | **SE** | **Lower CI** | **Upper CI** | ***P* (diff 0)** | ***P* (diff 1)** |
| UKB-PTSD | PGC1.5-PTSD | 0.5738 | 0.1155 | 0.3474 | 0.8001 | **6.83Ex10-7** | **0.0002** |
| UKB-PTSD | PGC2-PTSD | 0.8016 | 0.0703 | 0.6638 | 0.9394 | **3.80x10-30** | **0.005** |
| UKB-PTSD | MVP-PTSD | 0.6702 | 0.0707 | 0.5316 | 0.8088 | **2.56x10-21** | **3.10x10-6** |
| PGC1.5-PTSD | PGC2-PTSD | 0.9853 | 0.1481 | 0.6950 | 1.2756 | **2.85x10-11** | 0.92 |
| PGC1.5-PTSD | MVP-PTSD | 1.0053 | 0.1669 | 0.6782 | 1.3324 | **1.71x10-9** | 0.97 |
| PGC2-PTSD | MVP-PTSD | 0.9810 | 0.1105 | 0.7644 | 1.1976 | **6.96x10-19** | 0.86 |

Supplementary Table 4: High Definition Likelihood (HDL) genetic correlation estimates (*r*g), standard errors (SE), and 95% confidence intervals (CI) between the four major depressive disorder (MDD) categories: 1) MDD with reported trauma, 2) MDD without reported trauma, 3) recurrent MDD, 4) single-episode MDD. *P* (diff 0) refers to *p*-value to test whether the *r*g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the *r*g estimate differs from 1. Genetic correlations were considered significant if they surpassed the Bonferroni adjusted threshold (*p*<0.008) to correct for the 6 tests. Significant *p*-values are shown in bold.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Phenotype 1** | **Phenotype 2** | ***r*g** | **SE** | **Lower CI** | **Upper CI** | ***P* (diff 0)** | ***P* (diff 1)** |
| MDD with reported trauma | MDD without reported trauma | 0.6068 | 0.0727 | 0.4643 | 0.7493 | **7.11x10-17** | **6.35x10-8** |
| MDD with reported trauma | Recurrent MDD | 0.8120 | 0.0438 | 0.7262 | 0.8978 | **1.46x10-76** | **1.77x10-5** |
| MDD with reported trauma | Single-episode MDD | 0.8508 | 0.0629 | 0.7275 | 0.9741 | **1.13x10-41** | 0.02 |
| MDD without reported trauma | Recurrent MDD | 0.8686 | 0.0748 | 0.7222 | 1.0152 | **3.31x10-31** | 0.08 |
| MDD without reported trauma | Single-episode MDD | 0.9175 | 0.1013 | 0.7190 | 1.1160 | **1.35x10-19** | 0.42 |
| Recurrent MDD | Single-episode MDD | 0.9424 | 0.0596 | 0.8256 | 1.0592 | **2.49x10-56** | 0.33 |

Supplementary Table 5: genetic correlation results from the High Definition Likelihood (HDL) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). Each genetic correlation was compared in a pairwise fashion with all other genetic correlations in the set (i.e. for each PTSD phenotype). *r*g difference refers to the difference between the two genetic correlation estimates, SE refers to the standard error and *P* (diff 0) refers to *p*-value to test whether the *r*g difference differs significantly from 0. Differences between genetic correlations were considered statistically significant if they surpassed the Bonferroni adjusted threshold (p<0.008). Significant *p*-values are shown in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| HDL genetic correlation 1 | HDL genetic correlation 2 | *r*g difference | SE | *P* (diff 0) |
| UKB-PTSD and MDD with reported trauma | UKB-PTSD and MDD without reported trauma | 0.1339 | 0.0910 | 0.14 |
| UKB-PTSD and MDD with reported trauma | UKB-PTSD and recurrent MDD | -0.1094 | 0.0572 | 0.06 |
| UKB-PTSD and MDD with reported trauma | UKB-PTSD and single-episode MDD | -0.0426 | 0.0864 | 0.62 |
| UKB-PTSD and MDD without reported trauma | UKB-PTSD and recurrent MDD | -0.2433 | 0.0696 | **4.77x10-4** |
| UKB-PTSD and MDD without reported trauma | UKB-PTSD and single-episode MDD | -0.1765 | 0.0773 | 0.02 |
| UKB-PTSD and recurrent MDD | UKB-PTSD and single-episode MDD | 0.0668 | 0.0752 | 0.37 |
|  |  |  |  |  |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5-PTSD and MDD without reported trauma | 0.0679 | 0.0988 | 0.49 |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5-PTSD and recurrent MDD | -0.1417 | 0.0730 | 0.05 |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5-PTSD and single-episode MDD | -0.2040 | 0.1413 | 0.15 |
| PGC1.5-PTSD and MDD without reported trauma | PGC1.5-PTSD and recurrent MDD | -0.2096 | 0.0745 | **4.90x10-3** |
| PGC1.5-PTSD and MDD without reported trauma | PGC1.5-PTSD and single-episode MDD | -0.2719 | 0.1183 | 0.02 |
| PGC1.5-PTSD and recurrent MDD | PGC1.5-PTSD and single-episode MDD | -0.0623 | 0.1228 | 0.61 |
|  |  |  |  |  |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and MDD without reported trauma | 0.0988 | 0.1261 | 0.43 |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and recurrent MDD | -0.1418 | 0.0875 | 0.10 |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and single-episode MDD | -0.1650 | 0.1047 | 0.12 |
| PGC2-PTSD and MDD without reported trauma | PGC2**-**PTSD and recurrent MDD | -0.2406 | 0.0862 | **5.28x10-3** |
| PGC2-PTSD and MDD without reported trauma | PGC2**-**PTSD and single-episode MDD | -0.2638 | 0.1062 | 0.01 |
| PGC2-PTSD and recurrent MDD | PGC2**-**PTSD and single-episode MDD | 0.0232 | 0.1024 | 0.82 |
|  |  |  |  |  |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and MDD without reported trauma | 0.0538 | 0.1197 | 0.65 |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and recurrent MDD | -0.0203 | 0.0722 | 0.78 |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and single-episode MDD | -0.0894 | 0.1028 | 0.38 |
| MVP-PTSD and MDD without reported trauma | MVP**-**PTSD and recurrent MDD | -0.0741 | 0.0880 | 0.40 |
| MVP-PTSD and MDD without reported trauma | MVP**-**PTSD and single-episode MDD | -0.1432 | 0.1080 | 0.18 |
| MVP-PTSD and recurrent MDD | MVP**-**PTSD and single-episode MDD | -0.0691 | 0.1130 | 0.54 |

Supplementary Table 6: Linkage Disequilibrium Score Regression (LDSC) genetic correlation estimates (*r*g) and standard errors (SE) and 95% confidence intervals (CI) of 1) UK Biobank posttraumatic stress disorder (PTSD), 2) Psychiatric Genomics Consortium (PGC) 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD) with the four major depressive disorder (MDD) categories. *P* (diff 0) refers to *p*-value to test whether the *r*g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the *r*g estimate differs from 1. Genetic correlations were considered significant if they passed the Bonferroni adjusted threshold (*p*<0.0125). Significant *p*-values are shown in bold.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| PTSD Phenotype | MDD phenotype | *r*g | SE | Lower CI | Upper CI | *P* (diff 0) | *P* (diff 1) |
| UKB-PTSD | MDD with reported trauma | 0.6499 | 0.0832 | 0.4868 | 0.8130 | **5.70x10-15** | **2.58x10-5** |
| UKB-PTSD | MDD without reported trauma | 0.6538 | 0.1365 | 0.3863 | 0.9213 | **1.67x10-6** | **0.01** |
| UKB-PTSD | Recurrent MDD | 0.7815 | 0.0550 | 0.6737 | 0.8893 | **7.55x10-46** | **7.11x10-5** |
| UKB-PTSD | Single-episode MDD | 0.7158 | 0.1045 | 0.5110 | 0.9206 | **7.29x10-12** | **6.54x10-3** |
|  |  |  |  |  |  |  |  |
| PGC1.5-PTSD | MDD with reported trauma | 0.6976 | 0.2131 | 0.2799 | 1.1153 | **1.06x10-3** | 0.16 |
| PGC1.5-PTSD | MDD without reported trauma | 0.4514 | 0.2471 | -0.0329 | 0.9357 | **6.77x10-2** | 0.03 |
| PGC1.5-PTSD | Recurrent MDD | 0.7438 | 0.1545 | 0.4410 | 1.0466 | **1.48x10-6** | 0.10 |
| PGC1.5-PTSD | Single-episode MDD | 0.5743 | 0.1696 | 0.2418 | 0.9067 | **7.07x10-4** | **0.01** |
|  |  |  |  |  |  |  |  |
| PGC2-PTSD | MDD with reported trauma | 0.6263 | 0.1160 | 0.3989 | 0.8537 | **6.72x10-8** | **1.27x10-3** |
| PGC2-PTSD | MDD without reported trauma | 0.5461 | 0.1680 | 0.2168 | 0.8754 | **1.15x10-3** | **6.90x10-3** |
| PGC2-PTSD | Recurrent MDD | 0.7592 | 0.0880 | 0.5867 | 0.9317 | **6.47x10-18** | **6.21x10-3** |
| PGC2-PTSD | Single-episode MDD | 0.6597 | 0.1161 | 0.4321 | 0.8873 | **1.32x10-8** | **3.38x10-3** |
|  |  |  |  |  |  |  |  |
| MVP-PTSD | MDD with reported trauma | 0.3906 | 0.0766 | 0.2405 | 0.5407 | **3.40x10-7** | **1.78x10-15** |
| MVP-PTSD | MDD without reported trauma | 0.3593 | 0.1035 | 0.1564 | 0.5622 | **5.15x10-4** | **6.00x10-10** |
| MVP-PTSD | Recurrent MDD | 0.4812 | 0.0624 | 0.3589 | 0.6035 | **1.30x10-14** | **9.24x10-17** |
| MVP-PTSD | Single-episode MDD | 0.4392 | 0.0885 | 0.2657 | 0.6127 | **6.84x10-7** | **2.35x10-10** |

Supplementary Table 7: genetic correlation results from the Linkage Disequilibrium Score Regression (LDSC) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). *r*g difference refers to the difference between the two genetic correlation estimates, SE refers to the standard error and *P* (diff 0) refers to *p*-value to test whether the *r*g difference differs significantly from 0. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations in the set (i.e. for each PTSD phenotype). Differences between genetic correlations were considered statistically significant if they passed the Bonferroni adjusted threshold (p<0.008).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| LDSC genetic correlation 1 | LDSC genetic correlation 2 | *r*g difference | SE | *P* (diff 0) |
| UKB-PTSD and MDD with reported trauma | UKB**-**PTSD and MDD without reported trauma | -0.0039 | 0.1677 | 0.99 |
| UKB-PTSD and MDD with reported trauma | UKB**-**PTSD and recurrent MDD | -0.1316 | 0.0825 | 0.08 |
| UKB-PTSD and MDD with reported trauma | UKB**-**PTSD and single-episode MDD | -0.0659 | 0.1234 | 0.55 |
| UKB-PTSD and MDD without reported trauma | UKB**-**PTSD and recurrent MDD | -0.1277 | 0.1238 |  0.25 |
| UKB-PTSD and MDD without reported trauma | UKB**-**PTSD and single-episode MDD | -0.0620 | 0.1271 | 0.57 |
| UKB-PTSD and recurrent MDD | UKB**-**PTSD and single-episode MDD | 0.0657 | 0.1091 | 0.51 |
|  |  |  |  |  |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5**-**PTSD and MDD without reported trauma | 0.2462 | 0.2851 | 0.40 |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5**-**PTSD and recurrent MDD | -0.0462 | 0.1724 | 0.76 |
| PGC1.5-PTSD and MDD with reported trauma | PGC1.5**-**PTSD and single-episode MDD | 0.1233 | 0.2004 | 0.58 |
| PGC1.5-PTSD and MDD without reported trauma | PGC1.5**-**PTSD and recurrent MDD | -0.2924 | 0.2110 | 0.16 |
| PGC1.5-PTSD and MDD without reported trauma | PGC1.5**-**PTSD and single-episode MDD | -0.1229 | 0.2305 | 0.57 |
| PGC1.5-PTSD and recurrent MDD | PGC1.5**-**PTSD and single-episode MDD | 0.1695 | 0.1774 | 0.36 |
|  |  |  |  |  |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and MDD without reported trauma | 0.0802 | 0.1959 | 0.65 |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and recurrent MDD | -0.1329 | 0.1125 | 0.23 |
| PGC2-PTSD and MDD with reported trauma | PGC2**-**PTSD and single-episode MDD | -0.0334 | 0.1348 | 0.87 |
| PGC2-PTSD and MDD without reported trauma | PGC2**-**PTSD and recurrent MDD | -0.2131 | 0.1448 | 0.12 |
| PGC2-PTSD and MDD without reported trauma | PGC2**-**PTSD and single-episode MDD | -0.1136 | 0.1626 | 0.49 |
| PGC2-PTSD and recurrent MDD | PGC2**-**PTSD and single-episode MDD | 0.0995 | 0.1250 | 0.36 |
|  |  |  |  |  |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and MDD without reported trauma | 0.0313 | 0.1261 | 0.82 |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and recurrent MDD | -0.0906 | 0.0736 | 0.14 |
| MVP-PTSD and MDD with reported trauma | MVP**-**PTSD and single-episode MDD | -0.0486 | 0.0992 | 0.60 |
| MVP-PTSD and MDD without reported trauma | MVP**-**PTSD and recurrent MDD | -0.1219 | 0.0954 | 0.15 |
| MVP-PTSD and MDD without reported trauma | MVP**-**PTSD and single-episode MDD | -0.0799 | 0.0953 | 0.40 |
| MVP-PTSD and recurrent MDD | MVP**-**PTSD and single-episode MDD | 0.0420 | 0.0904 | 0.52 |

Supplementary Table 8: Results table from PRSice analysis. Polygenic Risk Scores (PRS) for postraumatic stress disorder (PTSD) were calculated based on the Million Veteran Program (MVP-PTSD) summary statistics. Table includes the best fitting *p*-value threshold used in the analyses (Threshold), the standardised beta coefficient from regression (Standardised beta) and standard error (SE), the odds ratio (OR) and 95% confidence intervals (CI), *p*-value from regression (*P*), the number of single nucleotide polymorphisms used in creating the risk scores (Number of SNPs) and empirical *p*-value accounting for testing at multiple thresholds (Empirical *P*). Regression coefficients were considered statistically significant if they surpassed the Bonferroni adjusted alpha (p<0.025). Significant *p*-values are shown in bold.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Regression | Threshold | Standardised beta | SE | OR (95% CI) | *P* | Number of SNPs | Empirical *P* |
| MDD with reported trauma vs. MDD without reported trauma | 0.4 | 0.04 | 0.014 |  1.04 (1.01 – 1.07) | 0.003 | 101111 | **0.02** |
| Recurrent MDD vs. single-episode MDD | 0.001 | -0.03 | 0.013 | 0.97 (0.95 – 0.99) |  0.01 | 646 | 0.08 |

**Supplementary References**

[**Allen NE, Sudlow C, Peakman T, Collins R, UK Biobank** (2014) UK biobank data: come and get it. *Science translational medicine* **6**, 224ed4.](http://paperpile.com/b/d3jKGH/DU6Pp)

[**Bellis MA, Hughes K, Leckenby N, Perkins C, Lowey H** (2014) National household survey of adverse childhood experiences and their relationship with resilience to health-harming behaviors in England. *BMC medicine* **12**, 72.](http://paperpile.com/b/d3jKGH/tGYdL)

[**Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, Sapareto E, Ruggiero J** (1994) Initial reliability and validity of a new retrospective measure of child abuse and neglect. *The American journal of psychiatry* **151**, 1132–1136.](http://paperpile.com/b/d3jKGH/Rodou)

[**Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R…Neale BM** (2015a) An atlas of genetic correlations across human diseases and traits. *Nature genetics* **47**, 1236–1241.](http://paperpile.com/b/d3jKGH/pjrcr)

[**Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium…Neale BM** (2015b) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291–295.](http://paperpile.com/b/d3jKGH/Kylad)

[**Burcusa SL, Iacono WG** (2007) Risk for recurrence in depression. *Clinical psychology review* **27**, 959–985.](http://paperpile.com/b/d3jKGH/frZy3)

[**Bush WS, Moore JH** (2012) Chapter 11: Genome-wide association studies. *PLoS computational biology* **8**, e1002822.](http://paperpile.com/b/d3jKGH/7ET8w)

[**Coleman JRI, Gaspar HA, Bryois J, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Breen G** (2019) The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biological psychiatry*.](http://paperpile.com/b/d3jKGH/8iWU4)

[**Coleman JRI, Peyrot WJ, Purves KL, Davis KAS, Rayner C, Choi SW…Breen G** (2020) Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Molecular psychiatry*](http://paperpile.com/b/bpS027/OovF)**25**, 1430-1446[.](http://paperpile.com/b/bpS027/OovF)

[**Davis KAS, Coleman JRI, Adams M, Allen N, Breen G, Cullen B…Hotopf M** (2020) Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych open* **6**, e18.](http://paperpile.com/b/d3jKGH/StIv)

[**Dudbridge F** (2013) Power and predictive accuracy of polygenic risk scores. *PLoS genetics* **9**, e1003348.](http://paperpile.com/b/d3jKGH/8DMtH)

[**Euesden J, Lewis CM, O’Reilly PF** (2015) PRSice: Polygenic Risk Score software. *Bioinformatics*  **31**, 1466–1468.](http://paperpile.com/b/d3jKGH/PRbBX)

[**Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J…O’Leary TJ** (2016) Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *Journal of clinical epidemiology* **70**, 214–223.](http://paperpile.com/b/d3jKGH/5qmKA)

[**Grabe HJ, Schulz A, Schmidt CO, Appel K, Driessen M, Wingenfeld K…Freyberger HJ** (2012) [A brief instrument for the assessment of childhood abuse and neglect: the childhood trauma screener (CTS)]. *Psychiatrische Praxis* **39**, 109–115.](http://paperpile.com/b/d3jKGH/yggKH)

[**Lee SH, Wray NR, Goddard ME, Visscher PM** (2011) Estimating missing heritability for disease from genome-wide association studies. *American journal of human genetics* **88**, 294–305.](http://paperpile.com/b/d3jKGH/nJmyE)

[**Nievergelt CM, Maihofer AX, Klengel T, Atkinson EG, Chen C-Y, Choi KW…Koenen KC** (2019) International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature communications* **10**, 4558.](http://paperpile.com/b/bpS027/yEsy)

[**Ning Z, Pawitan Y, Shen X** (2020) High-definition likelihood inference of genetic correlations across human complex traits. *Nature genetics* **52**, 859–864.](http://paperpile.com/b/d3jKGH/rx2BZ)

[**Stein MB, Levey DF, Cheng Z, Wendt FR, Harrington K, Cho K…Gelernter J** (2020) Genomic Characterization of Posttraumatic Stress Disorder in a Large US Military Veteran Sample. *BioRxiv*.](http://paperpile.com/b/bpS027/AJxq)

[**Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J…Collins R** (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* **12**, e1001779.](http://paperpile.com/b/bpS027/ayECb)