Isolating transdiagnostic effects reveals specific genetic profiles in psychiatric disorders.

Engin Keser1,5, Wangjingyi Liao2,5, Andrea G. Allegrini3, Kaili Rimfeld1,4, Thalia C. Eley1, Robert Plomin1,6 and Margherita Malanchini1,2,6

1 Social, Genetic and Developmental Psychiatry Centre, King's College London

2 School of Biological and Behavioural Sciences, Queen Mary University of London

3 Department of Psychology and Language Science, University College London

4 Department of Psychology, Royal Holloway University of London

5 These authors contributed equally

6 These authors jointly supervised the work

Correspondence: [Robert.plomin@kcl.ac.uk](mailto:Robert.plomin@kcl.ac.uk); [m.malanchini@qmul.ac.uk](mailto:m.malanchini@qmul.ac.uk)

# Abstract

Evidence indicates a great degree of genetic overlap between psychiatric diagnoses. Accounting for these transdiagnostic effects can sharpen research on disorder-specific genetic architecture. Here we isolate genetic effects that are shared across 11 major psychiatric disorders (p factor) to gain further insight into genetic specificity and comorbidity over and above that contributed by the p factor, unique to each psychiatric disorder. After adjusting for transdiagnostic genetic effects, we examined genetic correlations among psychiatric traits as well as relationships with other biobehavioural traits. The landscape of genetic associations between pairs of psychiatric disorders changed substantially, and their genetic correlations with biobehavioural traits showed greater specificity. Isolating transdiagnostic genetic effects across major psychiatric disorders provides a nuanced understanding of disorder-specific genetic architecture and genetic comorbidity, and may help guide diagnostic nomenclature and treatment research.

# Introduction

Genetic studies have challenged the current classification of psychiatric disorders as distinct categorical diagnoses by revealing overlaps in their genetic architectures 1. For instance, the first genome-wide association study (GWAS) of schizophrenia uncovered shared genetic loci between schizophrenia and bipolar disorder 2 despite their distinct categorization in diagnostic manuals (e.g., DSM-IV 3). Further research using Linkage Disequilibrium Score Regression (LDSC 4) highlighted significant genetic correlations across most psychiatric diagnoses 5-8, unlike other neurological disorders such as Parkinson’s and Alzheimer's disease, which remained genetically distinct 9. A recent LDSC analysis of 11 major psychiatric diagnoses found that the genetic correlation between schizophrenia and bipolar disorder was 0.68, and the average of 55 genetic correlations between 11 diagnoses was 0.28 10.

The positive genetic manifold between diagnoses is consistent with the idea of a general factor of psychopathology 11, called *p* 12. The p factor describes the propensity to developing all forms of psychopathologies 13, and reflects the comorbidities between psychiatric conditions that have been observed concurrently 14, across the lifespan 15,16, and even across generations 13,17. A p factor has also emerged from recent genetic and genomic studies 8,10,18. Shared genetic effects across different disorder dimensions were found to be stable over development, even when considering different measures and reporters 18. Capturing what cuts across diagnostic categories (i.e., a transdiagnostic approach) was found to be more effective in predicting functional and life outcomes than individual diagnoses 19.

Therefore, isolating p from major psychiatric disorders could better capture the specific genetic effects associated with each disorder and provide a more precise understanding of disorder-specific biology. Previous research that isolated transdiagnostic effects to investigate specificity revealed novel genetic profiles and biological pathways in neurodevelopmental disorders 20 and alcohol use disorder 21. In this study, we applied Genomic Structural Equation Modelling (Genomic SEM) 22 to isolate transdiagnostic genetic effects across 11 psychiatric disorders from genetic effects specific to each psychiatric condition. As shorthand, we refer to these disorder-specific genetic effects as *non-p* to describe residual genetic variance independent of genomically identified p. We used summary statistics from these non-p GWA analyses to estimate SNP heritability, genetic correlations between the disorders beyond transdiagnostic effects, and their correlations with external biobehavioural traits. We hypothesized that non-p traits will provide us with novel insight into the genetic architecture of psychiatric disorders and their correlations with other disorders and traits, consequently informing research and practice into diagnostics and treatment.

# Methods

The article is accompanied by Supplementary Information and the study followed a [preregistered analysis plan](https://osf.io/nq2wy). Deviations from preregistered analyses are described in Supplementary Note 1.

## GWAS summary statistics

We used the most recent publicly available summary statistics from GWA studies of 11 major psychiatric disorders. Detailed information about the GWAS summary statistics, sample sizes and availability are provided in Supplementary Table 1. All summary statistics are based on samples of European ancestry GWAS only. For details on the analysis protocol, we refer to the original publications.

## Genomic SEM

Genomic SEM 22 is a statistical framework that can model the shared and unique genetic architecture of complex traits by applying structural equation modelling principles to GWAS summary statistics. Genomic SEM is unbiased by sample overlap and imbalanced sample sizes 22. Here, we used Genomic SEM to perform multivariate GWAS analysis of 11 major psychiatric disorders (Figure 1) in order to capture transdiagnostic effects across all 11 disorders and isolate genetic variance in each psychiatric disorder beyond that captured by the p factor.

Genomic SEM uses multivariable LD score regression to estimate the genetic covariance matrix and sampling covariance matrix. We applied quality control filters for this step using the defaults in Genomic SEM, including restricting SNPs to those present in HapMap3 with a minor allele frequency > 1% and information score > 0.9. The LD weights used for LDSC were calculated using the European subsample of the 1000 Genomes phase 3 project; excluding the major histocompatibility complex (MHC) due to complex LD structures in this region that can bias estimates. When calculating the liability scale heritability estimates for the uncorrected psychiatric disorders, we used the sum of effective sample sizes, and a sample prevalence of 0.5 to reflect that the corrected sample size already accounts for sample ascertainment. See Supplementary Notes for more details about processing summary statistics in Genomic SEM and calculating effective sample sizes. After the quality control steps, 3,746,806 SNPs were present across all 11 disorders.

## Independent SNPs, Genes, and Enrichment and Pathway Analysis (MAGMA)

To identify independent hits from the 11 non-p GWAS, we applied a pruning approach using a window of 250 kb and an linkage disequilibrium (LD) threshold of r2 < 0.1. This was done using the LD clumping function in Plink v1.9 23, with LD statistics obtained from the 1000 Genome Project. Independent significant SNPs were considered novel if they were not present in the original disorder GWAS.

We used MAGMA within the FUMA framework (v1.5.6) 24 to map SNPs to genes based on their position (within 35kb upstream and 10kb downstream of each gene). We performed genome-wide gene-based association tests using MAGMA. The gene-based test combines results from multiple SNPs within a gene to assess the association between the gene and the disorder, while accounting for LD between SNPs. LD information was obtained from the 1000 Genomes Phase 3 EUR reference panel, and Bonferroni correction was applied to identify genes with genome-wide significance.

We used MAGMA to conduct tissue-specific gene-set analysis and gene property analysis. Gene-set analyses assessed whether genes within an annotated set exhibited stronger associations with the disorder compared to other genes. Meanwhile, the tissue specificity test examined the relationship between tissue-specific gene expression profiles and disorder-gene associations. The gene-set analyses were performed using curated gene sets and Gene Ontology (GO) terms obtained from the Molecular Signatures Database v2023.1Hs. For the MAGMA gene property analysis, tissue expression profiles were obtained from GTEx v8 (comprising 54 tissue types) and BrainSpan (brain samples at 11 general developmental stages), available in FUMA. Gene sets and tissue types were considered significant at a Bonferroni-corrected p-value threshold of <0.05.

## Genetic correlations

We used LDSC 4 to compute genetic correlations between the 11 psychiatric disorders after accounting for transdiagostic effects, as well as between each of the 11 psychiatric disorders –uncorrected and corrected for p– and 34 external traits. The LD scores used were computed using 1,215,002 SNPs present in the HapMap 3 reference panel, excluding the MHC region on chromosome 6. The differences between genetic correlations with external traits before and after partialling out genetic effects associated with the p factor were assessed using a two-stage method described in Coleman et al. (2020) 25. In this method, first, differences in genetic correlations were assessed using a two sample z-test, and significant differences (p < 0.05) were then compared using a block-jackknife correction. The results using the jackknife were then further corrected using the Benjamini-Hochberg False Discovery Rate (FDR) method to account for multiple testing.

# Results

## Isolating transdiagnostic genetic signal from 11 major psychiatric disorders

We used Genomic SEM to construct a genomic p factor using the most recent publicly available summary statistics from GWAS of 11 major psychiatric disorders (see Supplementary Table 1). Genetic correlations are presented in Supplementary Table 2. We found a positive manifold of genetic correlation (rG) among most disorders, with a mean rG of 0.29. Estimates ranged between -0.11 for the rG between obsessive compulsive disorder (OCD) and attention deficit hyperactivity disorder (ADHD) and 0.90 between anxiety and major depressive disorder (MDD).

To capture these transdiagnostic genetic effects across all 11 disorders, we fitted a common factor model to the genetic covariance matrix. In this model, all disorders loaded on a single common factor (i.e., the p-factor; top half of Fig. 1A). This same model also allowed us to capture residual genetic variance that was associated with each disorder independent of transdiagnostic effects (i.e., non-p; bottom half of Fig. 1A). We simultaneously ran GWAS on p and residual variance in each psychiatric disorder. Figure 1A provides a diagram of the GWAS that we ran to capture genetic variance in PTSD after accounting for transdiagnostic genetic effects. We repeated this procedure 11 times to isolate transdiagnostic genetic effects from each of the 11 major psychiatric disorders.

After accounting for the transdiagnostic genetic effects associated with p, we identified independent significant hits for seven out of eleven GWASs (see Method). The largest number of independent hits was observed for the non-p GWAS of schizophrenia (118, 27 of which were novel SNP associations that had not emerged as significant in the original GWAS), followed by BIP (22, 13 novel SNP associations), MDD (14, 11 novel SNP associations), ADHD (12, 8 novel SNP associations), AN (10, 8 novel SNP associations), ALCH (3, 1 novel SNP associations) and ASD (2). Some of these novel SNP associations had been uncovered by GWAS of psychiatric disorders other than the 11 included in our model, or by GWAS of other traits (not psychiatric disorders), while others had not been reported in the previous literature. For example, focusing on schizophrenia, 7 out of the 27 novel SNP associations had been uncovered by previous genomic studies on schizophrenia 26-29, 6 had been reported as SNPs associated with physiological or psychological traits (i.e.; body mass index (BMI) 30,31 and intelligence 32, and 14 SNP associations had not been reported in the extant literature. Details regarding the novel SNP uncovered for the other six non-p GWASs are presented in the Supplementary Tables 3-9.

We also used MAGMA33 in FUMA to evaluate the genetic effects of the 11 non-p GWASs (see Methods). The gene-level and gene-set analyses were performed to identify biological pathways linked to genes associated with each major psychiatric disorder before and after accounting for p, and to analyse tissue type enrichment. All results reported are Bonferroni's corrected to reduce multiple comparison problems. We found that after accounting for transdiagnostic genetic effects, 316 genes were associated with SCZ, 63 with BIP, 44 with MDD, 29 with ADHD, 37 with AN, 5 with ASD, and 1 with ALCH. Only SCZ, BIP and MDD showed significant enriched gene sets after accounting for p. SCZ non-p showed the most enrichments. It was enriched in six gene-sets related to neuron system (Supplementary Table 11). Other results are reported in the Supplementary Tables 10-19.

We tested whether common variants in genes specifically expressed in 53 Genotype-Tissue Expression (GTEx) tissues were enriched in their effects on psychiatric disorders (SCZ, BIP, MDD, ADHD, ALCH, ASD, AN) after accounting for transdiagnostic effects (See Methods). Genes predominantly expressed in the brain cortex and other brain-specific tissues were enriched in MDD, BIP, SCZ, ADHD, ASD and AN (Supplementary Figures 13-25, and Supplementary Table 12-47). Enrichment patterns were overall consistent between psychiatric disorders before and after removing transdiagnostic signals, but there were some exceptions. First, brain development stages enrichment results showed that, for SCZ, the early-to-late prenatal stages were enriched before accounting for p, but no longer enriched after removing transdiagnostic effects. Late infancy remained the most enriched developmental stage for SCZ, showing the strongest associations (Supplementary Figure 17, and Supplementary Table 21, 23). Second, for ASD, tissue property enrichment analyses showed that ten brain regions were enriched, with the substantia nigra showing the strongest signal, after accounting for the transdiagnostic effects, compared to four brain regions before (Supplementary Figure 21, and Supplementary Table 44, 46).



**Figure 1.** **Isolating transdiagnostic genetic effects from 11 major psychiatric disorders.** **A.** ***Standardized results for a common factor model of genomic p***. Each square indicates observed variables (i.e., the summary statistics for each of the 11 major psychiatric disorders) and circles represent latent variables that are statistically inferred from the data (i.e., genomic p-factor). One-headed arrows are standardized factor loadings, representing regression relations with the arrow pointing from the predictor variable to the outcome variable. Covariance relationships between variables are represented as two-headed arrows linking the variables. Residual variances of a variable are represented as a two-headed arrow connecting the variable to itself. SEs are shown in parentheses. The red arrows linking the SNP to both the p-factor and PTSD provide an example of the model used to partition genetic variance associated with transdiagnostic effects from the genetic variance specific to each disorder. We ran the model 11 times, isolating transdiagnostic effects from each psychiatric condition at a time. ANX = anxiety disorder; MDD = major depressive disorder; PTSD = post-traumatic stress disorder; BIP = bipolar disorder; SCZ = schizophrenia; ADHD = attention-deficit hyperactivity disorder; ASD = autism spectrum disorder; ALCH = problematic alcohol use; OCD = obsessive-compulsive disorder; AN = anorexia nervosa; TS = Tourette syndrome; p = general psychopathology factor. **B and C.** ***SNP-based heritability******estimates including and after accounting for transdiagnostic effects.*** SNP heritabilities were calculated before (on the liability scale) and after (on the observed scale) removing genetic effects shared with the p-factor. Error bars indicate standard errors.

## SNP heritabilities for disorders independent of p

We used LDSC to estimate SNP-based heritability (h2) for the 11 disorders before and after partialing out the genetic variance associated with p. Figures 1B and 1C, and Supplementary Table 48 show the SNP-based h2 estimates for the 11 disorders (on the liability scale; Figure 1B) and for the residual variance in each psychiatric disorder after accounting for the genetic effects associated with p (on the observed scale; Figure 1C). In general, the pattern of SNP h2 remained similar before and after controlling for p. For example, the highest SNP h2 were observed for OCD and TS (.49 and .49, respectively) and the lowest estimates for ALCH and PTSD (.02 and .04).

***Removing transdiagnostic genetic effects significantly changed genetic relationships between psychiatric disorders.***

Figure 2 compares genetic correlations between each disorder and the other 10 disorders before and after removing the genetic variance each has in common with p (see Supplementary Figure S2 for the full genetic correlation matrices as heat maps and Supplementary Table 2 and 4 for the genetic correlation estimates and confidence intervals.) It can be seen from Figure 2 that genetic correlations were lower after accounting for p. The average genetic correlation dropped from 0.29 to -0.05 after residualizing p. This reduction in genetic correlation suggests greater specificity in that the pervasive contribution of p to the positive manifold of genetic correlations is reduced.

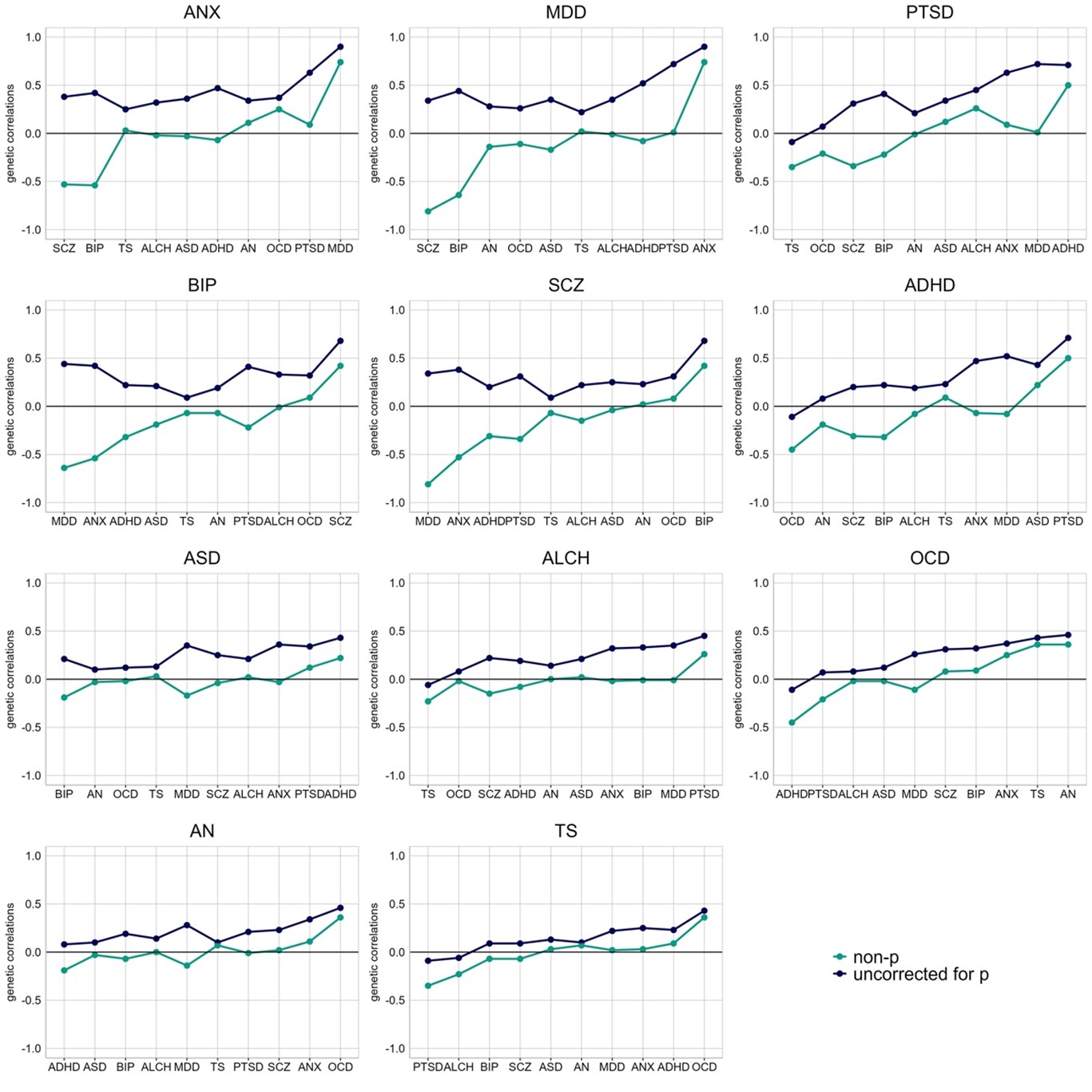
The major finding visualized in Figure 2 is that the reductions in transdiagnostic correlations are not uniform – that is, the blue and green lines are not parallel. For example, consider the genetic correlations between ADHD corrected and uncorrected for p with PTSD and MDD. Uncorrected for p, the genetic correlation between ADHD and PTSD is very high, 0.72. When ADHD and PTSD are corrected for p, the genetic correlation declines to 0.50. This means that the substantial genetic correlation between ADHD and PTSD is only slightly due to p. In contrast, for MDD, the genetic correlation with ADHD drops from 0.52 to -0.08. In other words, the substantial genetic correlation between ADHD and MDD disappears when ADHD is corrected for p, suggesting that the ostensible genetic relationship between ADHD and MDD is completely mediated by p.

The unexpected changes are cases where the genetic correlations between pairs of disorders goes from positive to negative. For example, the genetic correlation between BIP and MDD uncorrected for p was substantially positive (0.44), but after controlling for p, the genetic correlation became negative (-0.64).

Like ADHD and PTSD, for some pairs of disorders, genetic correlations remained substantial after accounting for transdiagnostic effects. For example, moderate to strong genetic correlations between ANX and MDD (rG = 0.74, SE = 0.08) and between SCZ and BIP (rG = 0.42, SE = 0.04) could still be observed. For other pairs of disorders, like ADHD and MDD, genetic associations were not significant after accounting for transdiagnostic genetic effects. This could be observed most notably for the otherwise strong correlation between PTSD and ANX (rG = 0.63, SE = 0.08), which was reduced to 0.09 (SE = 0.17) after removing the genetic variance they shared with p. Similarly, the moderate genetic correlation between ASD and ANX (rG = 0.36, SE = 0.05) dropped to -0.03 (SE = 0.10).

Another pattern of change was observed for the genetic correlations between OCD and ADHD and between TS and PTSD. For these disorder pairs, associations that were previously negative, but small or not significant, remained negative, but their effect size increased substantially. The genetic correlation between OCD and ADHD increased from -0.11 (SE = 0.06) to -0.45 (SE = 0.09), and the genetic correlation between TS and PTSD increased from -0.09 (SE = 0.10) to -0.35 (SE = 0.19). The pronounced change emerged between MDD, BIP and SCZ, in which genetic correlations switched from positive to negative after removing transdiagnostic genetic effects. For MDD and SCZ, the genetic correlation changed from 0.34 (SE = .03) to -0.81 (SE = 0.05). The change was similar for the association between MDD and BIP, from 0.44 (SE = 0.03) to -0.64 (SE = 0.06).

The same pattern of genetic correlations was obtained when isolated transdiagnostic effects and obtained disorder-specific summary statistics using a different method that followed a two-step procedure: 1) we created a genomic p factor and 2) we used GWAS-by-subtraction 34 to separate genetic effects associated with the genomic p-factor constructed at Step 1 from the genetic effects associated with each psychiatric disorder (See Supplementary Note 5).



**Figure 2| Changes in the landscape of genetic correlations between psychiatric disorders after accounting for transdiagnostic effects.** For every psychiatric disorder, we presentgenetic correlations with the other 10 psychiatric conditions uncorrected for p (blue line) and the genetic correlations with the other psychiatric disorders after removing the genetic variance captured by the p factor (non-p; green line). Disorders are presented ordered by their loading on the p factor, starting from the highest-loading disorders in the top row (i.e., PTSD, MDD and ANX) to those that shared the least amount of genetic variance with the p factor in the bottom row (i.e., AN and OCD). Correlations were estimated using LDSC within Genomic SEM. ANX, anxiety disorder; MDD, major depressive disorder; PTSD, post-traumatic stress disorder; BIP, bipolar disorder; SCZ, schizophrenia; ADHD = attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; ALCH, problematic alcohol use; OCD, obsessive-compulsive disorder; AN, anorexia nervosa; TS, Tourette syndrome.

## Genetic architecture of the associations with external traits

In addition to removing transdiagnostic genetic effects from the relationships between psychiatric disorders, we also compared genetic correlations between each of the 11 psychiatric disorders – uncorrected and corrected for p – and 34 traits that are not psychiatric disorders. We focused on four broad categories of external traits: socio-demographic, anthropometric, health-related, and psychological traits. These correlations are shown in Figure 3 and 4 and Supplementary Figures 26 and 27, with details in Supplementary Table 50.

**Figure 3. Genetic correlations between 11 major psychiatric disorders and psychological traits.** The dots represent genetic correlations estimated using LDSC regression. Correlations with psychiatric disorders uncorrected for p (‘original’) are in blue, with psychiatric disorders corrected for p (‘non-p’) in green. Error bars represent 95% confidence intervals. Red asterisks indicate a statistically significant (FDR-corrected *P* < 0.05, two-tailed test) differences in the magnitude of the correlation with disorders uncorrected for p versus disorders corrected for p. Exact *P* values for all associations are reported in Supplementary Table 50. The FDR correction was applied based on all genetic correlations tested (including those reported in Supplementary Figures 26-27). Source GWASs are listed in Supplementary Table 51.

Figure 3 presents genetic correlations for the 11 psychiatric disorders uncorrected and corrected for p and psychological traits. Correlations between disorders uncorrected and corrected for p and most psychological traits differed significantly, as indicated by the red asterisks. In almost all cases, the genetic correlations were lower for the corrected than for the uncorrected psychiatric traits. Importantly, correlations with the same traits often appeared to be significantly different across major psychiatric disorders, such as subjective wellbeing, sensitivity to environmental stress and tiredness. For example, consider sensitivity to environmental stress. Uncorrected for p, genetic correlations were significant and substantial between sensitivity to environmental stress and major psychiatric disorders, but corrected for p, these correlations were significantly lower and sometimes negligible: SCZ (0.19 vs -0.02), MDD (0.53 vs 0.03), ASD (0.19 vs. -0.002) and ALCH (0.29 vs. 0.04). A similar pattern of results was observed for several external traits that often co-occur with most psychiatric disorders, such as subjective wellbeing, loneliness, tiredness and insomnia. Supplementary Figures 26-27 and Supplementary Table 50 present the genetic correlations between psychiatric disorders and all socio-demographic, anthropometric, health-related and psychological traits.

As shown in Figure 4, similar results emerged for health-related traits. Genetic correlations between disorders corrected for p were generally lower than those uncorrected for p, often significantly lower. For example, self-reported poor health showed significant reductions for most disorders (ADHD, ALCH, ANX, BIP, MDD, SCZ); number of sexual partners showed significant reductions for four disorders (ADHD, ALCH, BIP, SCZ); and cigarette per day yielded significant reductions for three disorders (ALCH, ANX, MDD; see Supplementary Table 50).

Though most of the changes in correlations were reductions, there were some exceptions. For example, after accounting for p, the genetic correlation between BIP and SCZ and educational attainment increased significantly. Other genetic correlations reversed after accounting for p. For example, the genetic correlation between BIP and loneliness changed from rG = 0.11 before accounting for p to rG = -0.26 after removing transdiagnostic effects, a similar change was also observed for the genetic correlations between BIP and Back Pain (from 0.08 to -0.15) and MDD and several cognitive traits (cognitive performance (from -0.10 to 0.18), executive functions (from -0.18 to 0.31) and noncognitive skills (from -0.09 to 0.22) and risk-taking behaviours (e.g., the correlation between MDD and risk tolerance changed from 0.14 to -0.29). Significant changes were also observed for the genetic correlations between ANX and risk-taking behaviour (from 0.14 to -0.26), ANX and risk tolerance (from 0.07 to -0.39), and SCZ and back pain (from 0.05 to -0.16).





**Figure 4. Genetic correlations between 11 major psychiatric disorders and health-related traits.** The dots represent genetic correlations estimated using LDSC regression. Correlations with psychiatric disorders uncorrected for p (‘original’) are in blue, with psychiatric disorders corrected for p (‘non-p’) in green. Error bars represent 95% confidence intervals. Red asterisks indicate a statistically significant (FDR-corrected *P* < 0.05, two-tailed test) differences in the magnitude of the correlation with disorders uncorrected for p versus disorders corrected for p. Exact *P* values for all associations are reported in Supplementary Table 50. The FDR correction was applied based on all genetic correlations tested (including those reported in Supplementary Figures 26-27).

# DISCUSSION

We isolated transdiagnostic genetic effects across 11 major psychiatric disorders from genetic effects specific to each psychiatric condition. We uncovered differences in the genetic architecture of each disorder, identifying novel genome-wide significant SNP associations for four psychiatric conditions —schizophrenia, bipolar disorder, ADHD and anorexia nervosa— and specific changes in tissue enrichment. The greatest shift we observed, however, was in the genetic relationships between disorders, which changed substantially after considering transdiagnostic effects. Genetic effects associated with individual disorders beyond p also showed a greater degree of specificity in their genetic correlations with external biobehavioral traits, particularly health-related and psychological traits. By separating genetic influences into a shared and specific components, we provide novel insight into the genetic architecture, biology and comorbidity between psychiatric conditions that can inform diagnostic nomenclature and treatment research.

While our non-p GWA analyses revealed fewer genome-wide significant SNP associations than the GWA analyses uncorrected for p, they uncovered several novel trait-specific hits. For example, of the 27 novel SNP associations with SCZ, one of the top hit (rs693906) has not been previously reported by GWASs before, but was an intronic SNP in a gene (SLC44A4) associated with schizophrenia in a Japanese sample 35. Such novel non-p SNPs associations might provide new leads to uncover biological pathways specific to schizophrenia. Our analyses mapping risk variants to gene sets and tissue expression revealed specific changes in tissue enrichment beyond transdiagnostic effects, these were almost exclusively observed for SCZ and BIP. SNP heritabilities for most disorders remained stable when accounting for transdiagnostic effects, with the notable exceptions of MDD and ANX, which exhibited significant reductions in genetic variance.

We were particularly interested in investigating the genetic correlations between psychiatric disorders. We found that removing transdiagnostic genetic effects significantly changed genetic relationships between psychiatric conditions. While some genetic correlations showed a reduction in magnitude, the pattern of genetic correlations between several disorders shifted unexpectedly, highlighting discrepancies between diagnostic nosology and genetic structure of psychiatric disorders. For some pairs of disorders, positive genetic correlations disappeared entirely after correction for p, such as the genetic correlation between ANX and PTSD, indicating that their overlap is likely entirely captured by what cuts across all psychiatric diagnoses. Other pairs of disorders remained correlated positively after accounting for p, such as the genetic correlation between SCZ and BIP and between ANX and MDD, which suggests that only part of their overlap is shared with the other disorders included in our transdiagnostic model. These differences in genetic correlations are not simply due to the disorders’ loading on the genomic p factor. For example, loadings on the genomic p factor are similar for PTSD (0.78) and MDD (0.82), but the p-corrected genetic correlation with ADHD is much reduced for MDD (from 0.52 to -0.08) but not for PTSD (0.72 to 0.50). These predictive profile differences reveal differences in disorder-specific genetic architecture when the positive manifold of transdiagnostic effects of genomic p is removed.

Other genetic correlations presented a different pattern of change. The genetic correlations between SCZ and MDD and between MDD and BIP switched from strongly positive to strongly negative after considering transdiagnostic effects. This negative correlation implies that, once transdiagnostic effects are accounted for, a genetic liability for one disorder conveys a lower genetic liability for the other disorder, this has the potential to inform diagnostic categorization and has far-reaching implications for future studies aimed at understanding biological causes and treatments for these disorders. Another major shift was observed for the slightly negative genetic correlation between OCD and ADHD, which became strongly negative after accounting for p. This suggests that their genetic overlap with p obscured the negative genetic relationship between these disorders. This finding is consistent with the current clinical perspective suggesting that OCD and ADHD lie at the opposite extremes of the impulsivity-compulsivity continuum 36,37.

Comparing genetic correlations between the 11 psychiatric disorders and external biobehavioural traits further indicated greater specificity for disorders corrected for p, especially for psychological and health-related traits. Though most of the changes in genetic correlations were reductions, there were some exceptions, with some correlations increasing and other genetic correlations reversing. Removing transdiagnostic genetic effects from psychiatric disorders, particularly MDD, ANX, SCZ and BIP, resulted in divergent patterns of genetic associations with external biobehavioural traits. In other words, genetic effects associated with each major psychiatric disorder also include overlapping genetic variance associated with these transdiagnostic traits, which mask their specificity.



Together, our findings highlight how isolating transdiagnostic genetic risk from major psychiatric disorders provides novel insight into disorder-specific genetic architecture and a more nuanced understanding of their comorbidities and co-occurrences with psychological and health-related traits. Consequently, our findings emphasize the significance of considering specificity as well as generality in psychiatric genetics. By demonstrating distinct genetic correlations and outcomes associated with psychiatric conditions independent of transdiagnostic effects, our findings pave the way for new avenues of research. One such application is the use of non-p summary statistics to create polygenic scores that index greater specificity in psychiatric disorders. For example, isolating transdiagnostic effects from psychiatric polygenic risk scores can lead to refined predictions of how biological risk for each disorder unfolds developmentally and can offer new avenues to investigate how genetic and environmental risk combine.

**Limitation**

Our findings need to be interpreted in the context of their limitations. The most important limitation is that our research begins with case-control GWAS based on traditional diagnoses perfused with transdiagnostic effects. This points to the need for GWA research using phenotypes that correspond more closely to the genetic architecture of psychopathology. The origin of psychiatric nosology is historical rather than empirical; progress depends on more empirically derived dimensional approaches such as HiTop 38 and Rdoc 39.

Second, other limitations of our study reflect broader challenges in GWA research. For example, the GWA studies that are the basis for this research are largely limited to individuals of European ancestry, so the results reported here might not generalize beyond this population 40,41. In addition, the contributing GWA studies are meta-analyses of different cohorts that may be subject to heterogeneity that cannot be fully quantified. Moreover, our research is limited by issues that could affect any GWA results, such as cross-trait assortative mating 42 and population stratification 43.

Two additional considerations are important when interpreting our results. While SEM is a powerful tool for modelling genetic overlap across traits, the resulting factors do not represent causal entities 44 and are sensitive to constructs included in the model 45. The p-factor therefore reflects patterns of genetic covariance among the eleven disorders as captured in the current GWA studies. As such, the p-factor should not be interpreted as a definitive index of general psychopathology, but rather as a data-driven summary of genetic covariance across the included disorders. Similarly, non-p PGSs were constructed by statistically removing variance associated with p, but this process is imperfect. Therefore, non-p PGSs should be interpreted as approximations of residual trait-specific liability, not as pure indicators of specificity. We suggest that p will be similarly valuable for understanding general genetic influences, and that isolating the transdiagnostic effects captured by p will be useful in sharpening research on specific genetic influences, particularly in the context of developmental psychopathology and clinical epidemiological studies. Genetic effects that are disorder-specific might inform future research into causes and consequences of psychiatric conditions applying causal designs including mendelian randomization and longitudinal models 46,47.

Second point is related to our choice of modelling p as a common factor, given our interest in capturing transdiagnostic genetic effects that could index shared genetic liability across all 11 major psychiatric disorders. Although alternative models have been proposed 10,17,48 different statistical approaches to modelling p were found to lead to similar estimates 17. Relatedly, we allowed our indicators to load feely onto the common factor, as such, some disorders (e.g., ANX, MDD and PTSD) contributed more than others to the general factor (e.g., OCD and AN). A model in which all indicators are restricted to contribute the same amount of variance to the general factor would likely have led to different results, although arguably it would have provided a poorer account of transdiagnostic effects in psychopathology. It should also be noted that, while general factor models can fit psychopathological data, alternative explanations have been proposed, most notably network models 49.

**Conclusion**

In conclusion, our results show that isolating transdiagnostic effects from major psychiatric disorders provides novel insight into disorder-specific genetic architecture a more precise understanding of the comorbidities and co-occurrences in psychopathology. Until better correspondence between psychiatric diagnoses and the genetic architecture of psychopathology is achieved, isolating p from diagnostic categories will sharpen genetic research by focusing on disorder-specific genetic effects.

# Acknowledgements

E.K. is supported by a PhD scholarship from the Economic and Social Research Council. W.L. is supported by a Chinese Scholarship Council PhD fellowship. MM is supported by a starting grant from the School of Biological and Behavioural Sciences at Queen Mary University of London. We would like to thank Jonathan Coleman for sharing the scripts used to apply block-jackknife correction, which were written by Héléna Gaspar and Christopher Hübel.

# Conflict of Interest

The authors declare no conflict of interest.

# Contributions

E.K., W.L., R.P. and M.M. conceived and designed the study.; E.K. and W.L. analyzed the data with contributions from M.M. and A.G.A.. E.K., W.L., R.P. and M.M. wrote the paper with contributions from K.R., A.G.A., and T.C.E. All authors contributed to the interpretation of data, provided critical feedback on manuscript drafts, and approved the final draft.

# Code availability

Code will be available at <https://github.com/CoDEresearchlab> upon publication. Ahead of publication we will make the code available upon request.

# References

1. Plomin, R. The next 10 years of behavioural genomic research. *JCPP Adv* **2**, e12112 (2022).

2. International Schizophrenia, C. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-52 (2009).

3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, (Washington, DC, 2000).

4. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).

5. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).

6. Cross-Disorder Group of the Psychiatric Genomics, C. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984-94 (2013).

7. Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address, p.m.h.e. & Cross-Disorder Group of the Psychiatric Genomics, C. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482 e11 (2019).

8. Selzam, S., Coleman, J.R.I., Caspi, A., Moffitt, T.E. & Plomin, R. A polygenic p factor for major psychiatric disorders. *Transl Psychiatry* **8**, 205 (2018).

9. Brainstorm, C. *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**(2018).

10. Grotzinger, A.D. *et al.* Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat Genet* **54**, 548-559 (2022).

11. Lahey, B.B. *et al.* Is there a general factor of prevalent psychopathology during adulthood? *J Abnorm Psychol* **121**, 971-7 (2012).

12. Caspi, A. & Moffitt, T.E. All for One and One for All: Mental Disorders in One Dimension. *Am J Psychiatry* **175**, 831-844 (2018).

13. Caspi, A. *et al.* The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clin Psychol Sci* **2**, 119-137 (2014).

14. de Jonge, P. *et al.* The cross-national structure of mental disorders: results from the World Mental Health Surveys. *Psychol Med* **48**, 2073-2084 (2018).

15. Caspi, A. *et al.* Longitudinal Assessment of Mental Health Disorders and Comorbidities Across 4 Decades Among Participants in the Dunedin Birth Cohort Study. *JAMA Netw Open* **3**, e203221 (2020).

16. Plana-Ripoll, O. *et al.* Exploring Comorbidity Within Mental Disorders Among a Danish National Population. *JAMA Psychiatry* **76**, 259-270 (2019).

17. Caspi, A., Houts, R.M., Fisher, H.L., Danese, A. & Moffitt, T.E. The general factor of psychopathology (p): Choosing among competing models and interpreting p. *Clin Psychol Sci* **12**, 53-82 (2024).

18. Allegrini, A.G. *et al.* The p factor: genetic analyses support a general dimension of psychopathology in childhood and adolescence. *J Child Psychol Psychiatry* **61**, 30-39 (2020).

19. Eaton, N.R. *et al.* A review of approaches and models in psychopathology conceptualization research. *Nature Reviews Psychology* **2**, 622-636 (2023).

20. Pettersson, E., Anckarsater, H., Gillberg, C. & Lichtenstein, P. Different neurodevelopmental symptoms have a common genetic etiology. *J Child Psychol Psychiatry* **54**, 1356-65 (2013).

21. Mallard, T.T. *et al.* Item-Level Genome-Wide Association Study of the Alcohol Use Disorders Identification Test in Three Population-Based Cohorts. *Am J Psychiatry* **179**, 58-70 (2022).

22. Grotzinger, A.D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513-525 (2019).

23. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).

24. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).

25. Coleman, J.R.I. *et al.* Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Mol Psychiatry* **25**, 1430-1446 (2020).

26. Goes, F.S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am J Med Genet B Neuropsychiatr Genet* **168**, 649-59 (2015).

27. Bhattacharya, A. *et al.* Isoform-level transcriptome-wide association uncovers extensive novel genetic risk mechanisms for neuropsychiatric disorders in the human brain. *medRxiv*, 2022.08.23.22279134 (2023).

28. Chen, J. *et al.* Shared Genetic Risk of Schizophrenia and Gray Matter Reduction in 6p22.1. *Schizophr Bull* **45**, 222-232 (2019).

29. Ikeda, M. *et al.* Genome-Wide Association Study Detected Novel Susceptibility Genes for Schizophrenia and Shared Trans-Populations/Diseases Genetic Effect. *Schizophr Bull* **45**, 824-834 (2019).

30. Christakoudi, S., Riboli, E., Evangelou, E. & Tsilidis, K.K. Associations of body shape index (ABSI) and hip index with liver, metabolic, and inflammatory biomarkers in the UK Biobank cohort. *Sci Rep* **12**, 8812 (2022).

31. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* **53**, 1415-1424 (2021).

32. Hill, W.D. *et al.* A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry* **24**, 169-181 (2019).

33. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).

34. Demange, P.A. *et al.* Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat Genet* **53**, 35-44 (2021).

35. Yamada, K. *et al.* Population-dependent contribution of the major histocompatibility complex region to schizophrenia susceptibility. *Schizophr Res* **168**, 444-9 (2015).

36. Abramovitch, A., Dar, R., Mittelman, A. & Wilhelm, S. Comorbidity Between Attention Deficit/Hyperactivity Disorder and Obsessive-Compulsive Disorder Across the Lifespan: A Systematic and Critical Review. *Harv Rev Psychiatry* **23**, 245-62 (2015).

37. Hollander, E. Obsessive-compulsive disorder and spectrum across the life span. *Int J Psychiatry Clin Pract* **9**, 79-86 (2005).

38. Kotov, R. *et al.* The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J Abnorm Psychol* **126**, 454-477 (2017).

39. Insel, T.R. & Cuthbert, B.N. Endophenotypes: bridging genomic complexity and disorder heterogeneity. *Biol Psychiatry* **66**, 988-9 (2009).

40. Turley, P. *et al.* Multi-Ancestry Meta-Analysis yields novel genetic discoveries and ancestry-specific associations. *bioRxiv*, 2021.04.23.441003 (2021).

41. Ding, Y. *et al.* Polygenic scoring accuracy varies across the genetic ancestry continuum. *Nature* **618**, 774-781 (2023).

42. Border, R. *et al.* Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* **378**, 754-761 (2022).

43. Novembre, J. *et al.* Genes mirror geography within Europe. *Nature* **456**, 98-101 (2008).

44. van Loo, H.M., Aggen, S.H. & Kendler, K.S. The structure of the symptoms of major depression: Factor analysis of a lifetime worst episode of depressive symptoms in a large general population sample. *J Affect Disord* **307**, 115-124 (2022).

45. Watts, A.L., Lane, S.P., Bonifay, W., Steinley, D. & Meyer, F.A.C. Building theories on top of, and not independent of, statistical models: The case of the p-factor. *Psychol Inq* **31**, 310-320 (2020).

46. Pingault, J.-B., Cecil, C.A.M., Murray, J., Munafò, M.R. & Viding, E. Causal Inference in Psychopathology: A Systematic Review of Mendelian Randomisation Studies Aiming to Identify Environmental Risk Factors for Psychopathology. *Psychopathology Review* **a4**, 4-25 (2016).

47. Allegrini, A.G., Baldwin, J.R., Barkhuizen, W. & Pingault, J.B. Research Review: A guide to computing and implementing polygenic scores in developmental research. *J Child Psychol Psychiatry* **63**, 1111-1124 (2022).

48. Watts, A.L., Poore, H.E. & Waldman, I.D. Riskier Tests of the Validity of the Bifactor Model of Psychopathology. *Clinical Psychological Science* **7**, 1285-1303 (2019).

49. van Bork, R., Epskamp, S., Rhemtulla, M., Borsboom, D. & van der Maas, H.L.J. What is the p-factor of psychopathology? Some risks of general factor modeling. *Theory & Psychology* **27**, 759-773 (2017).