

Common genetic variants are associated with increased likelihood of some co-occurring mental health conditions among autistic individuals

AUTHORS: Adeniran Okewole¹, Vincent-Raphael Bourque^{1,2,3}, Mahmoud Koko⁴, Guillaume Huguet², Anders D Borglum^{5,6,7}, Jakob Grove^{5,6,7,8}, Sebastien Jacquemont^{2,9}, Simon Baron-Cohen^{1,10}, and Varun Warriar^{1,10}

AFFILIATIONS: 1. Autism Research Centre, Department of Psychiatry, Cambridge University, United Kingdom

2. CHU Sainte-Justine Pediatric Hospital Research Center, Montréal, Canada

3. Department of Psychiatry, McGill University, Montréal, Canada

4. Wellcome Sanger Institute, Hinxton, United Kingdom

5. The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark

6. Center for Genomics and Personalized Medicine (CGPM), Aarhus University, Aarhus, Denmark

7. Department of Biomedicine, Aarhus University, Aarhus, Denmark

8. Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark

9. Department of Pediatrics, Université de Montréal, Montréal, Canada

10. Department of Psychology, University of Cambridge

Correspondence to Adeniran Okewole (ao508@medschl.cam.ac.uk) or Varun Warriar (vw260@cam.ac.uk)

Key Points:

Question: Do genetic variants contribute to co-occurring mental health conditions and latent factors in autistic individuals?

Findings: In this cross-sectional study of 17,582 autistic individuals with and without co-occurring conditions, we found significant single nucleotide polymorphism (SNP) heritability for co-occurring ADHD, Disruptive Behaviour Disorders (DBD), and mental-health latent factors. Co-occurring ADHD, DBD, anxiety, and all mental health factors among autistic individuals had moderate genetic correlations with corresponding case-control GWAS. We found no evidence linking common genetic variants linked to autism and rare genetic variants with increased likelihood for mental health conditions among autistic individuals.

Meaning: Among autistic individuals, the genetic correlates of co-occurring mental health conditions are distinct from that of autism, suggesting that additional genetic factors contribute to the development of these conditions among autistic individuals.

ABSTRACT

IMPORTANCE: Autism frequently co-occurs with other mental health conditions. In the general population, these co-occurring mental health conditions are highly heritable and genetically correlated; however, the genetic architecture of co-occurring mental health conditions among autistic individuals is unclear.

OBJECTIVE: To investigate the relationship between common and rare genetic variation and co-occurring mental health conditions and latent factors among autistic individuals.

DESIGN: Cross-sectional

SETTING: The study was conducted with the Simons Foundation Powering Autism Research (SPARK) dataset, V9 release (12 December 2022).

PARTICIPANTS: Phenotypic data exploration and factor analyses was conducted in 74,204 autistic individuals, and genetic analyses were conducted in a maximum of 17,582 individuals with genetic data.

MAIN OUTCOMES AND MEASURES: Genetic analysis was limited to those probands included in the SPARK iWES1 dataset [n=17,582]. SNP heritability estimates and genetic correlations were computed for three factor scores and six diagnostic categories (attention deficit hyperactivity disorder or ADHD, bipolar disorder, depression, schizophrenia, anxiety disorder and disruptive behaviour disorders or DBD) using bivariate GCTA-GREML and LDSC . Polygenic scores were generated using summary statistics from the most recent genome wide association studies

(GWAS) for five traits and six conditions. Associations with factor scores and categorical diagnoses were tested separately for polygenic scores (PGS), *de novo* variants (DNVs) and copy number variants (CNVs).

RESULTS: 56% of autistic individuals presented a co-occurring psychiatric condition. Confirmatory factor analysis identified three minimally correlated factors: a behavioural factor, cothymic factor, and a 'Kraepelin' or thought disorder factor, with SNP heritabilities ranging from 0.21 (s.e. 0.02) for behavioural to 0.09 (s.e. 0.03) for cothymic factor. Among conditions, moderate and significant SNP heritabilities were observed for ADHD (0.18, s.e. = 0.04) and DBD (0.52, s.e. = 0.08). Moderate positive genetic correlations were found between co-occurring ADHD, DBD, anxiety and the three factors in autistic individuals and corresponding conditions in external population cohorts. PGS for ADHD, depression, and educational attainment were significantly associated with all mental health factors and some of the conditions tested. We found no evidence for an association between common variants for autism, rare CNVs, and DNVs in highly constrained genes with increased likelihood of mental health phenotypes among autistic individuals.

CONCLUSION AND RELEVANCE: Among autistic individuals, some mental health conditions and all mental health factors are heritable, but have a distinct genetic architecture from autism itself.

1. Introduction

Autism commonly co-occurs with several mental health conditions, with resultant impairment in quality of life^{1,2}. Approximately 70% of autistic children meet the criteria for at least one psychiatric condition³, and between 54 - 76% of autistic adults have at least one psychiatric diagnosis^{4,5}. Understanding why autistic individuals have higher prevalence of mental health conditions is an important immediate issue to be addressed.

Socio-contextual determinants and adverse life experiences are two significant factors associated with increased mental health conditions among autistic individuals¹. A third factor is genetics. Autism is heritable^{14,15}, and is associated with the additive effects of genetic variants across the allelic frequency spectrum^{16–21}. In external population cohorts, mental health conditions are partly heritable and genetically correlated^{22–24}. Consequently, it is conceivable that mental health conditions among autistic individuals are associated with genetic variants. However, no study to our knowledge has examined this.

At least two hypotheses can explain the link between genetics and mental health conditions among autistic individuals. Common genetic variants linked to autism are moderately positively genetically correlated with ADHD, depression, self-harm behaviours and modestly with schizophrenia, bipolar disorder and eating disorders^{14,25–27}. Similarly, rare genetic variants associated with autism overlap with ADHD²⁸ and schizophrenia^{29,30}. Consequently, one hypothesis is that co-occurring mental health conditions among autistic individuals are heritable and genetically correlated with genetic variants associated with autism. In other words, variation in genetic predispositions to autism among autistic individuals is associated with their mental health profiles. A second hypothesis is that people are more likely to receive a diagnosis of autism if they have additional mental health conditions and corresponding genetic variants. This would suggest that genetic influences distinct from those linked to autism are associated with elevated mental health conditions among autistic individuals.

Mental health conditions often co-occur³¹, even among autistic individuals, possibly because of shared features across these mental health conditions. Latent factor models have demonstrated transdiagnostic utility in the general population³², but their applicability to the autistic population is less well established. Studying the associations between latent mental health factors and genetics may help address potential underdiagnosis of mental health conditions among autistic individuals.

Here, we investigated the genetic correlates (common and rare genetic variants) of co-occurring mental health profiles (both categorical conditions and latent factors) in 17,582 autistic individuals from the Simons Foundation Powering Autism Research (SPARK) cohort³³. We find

significant SNP heritability for some categorical conditions and all latent factors. We do not find an association between common genetic variants for autism or rare genetic variants and mental health conditions. In contrast, we identify moderate to high genetic correlations between some mental health phenotypes and their corresponding phenotypes in the external population cohorts. Finally, within-family based analyses demonstrate direct effects of polygenic scores on some mental health phenotypes without evidence for indirect genetic effects through the familial environment.

2. Materials and Methods

Note on terminology

We use “autistic individuals” in line with preferences of the autistic community. We use preferentially neutral terminology, i.e. “likelihood” or “probability”, when referring to autism and co-occurring conditions. Sex refers to sex assigned at birth.

2.1. Description of dataset

Participants in this study included individuals with an autism diagnosis in the Simons Foundation Powering Autism Research (SPARK) dataset³³ (**Supplementary Figure 1**). Phenotype data was available in the V9 release (12 December 2022) and included 123,444 autistic and non-autistic individuals. Phenotype analysis involved 74,204 autistic individuals for whom core descriptive information was available. Genetic data was obtained from the SPARK iWES v2 (January 2023) release, which included 105,838 autistic and non-autistic individuals. These included 69,758 autistic and non-autistic individuals previously in iWES v1 (WES1-4), who were genotyped using array genotyping. To ensure homogeneity in genotyping across all participants and to allow for joint common and rare variant exploration, genetic analysis was limited to 17,582 participants in WES1 and passed quality control.

2.2. Overview of phenotypes, data exploration and confirmatory factor analysis

We focussed on the profiles of mental health conditions in 74,204 autistic individuals (**Supplementary Figure 2**). These include ADHD, schizophrenia, bipolar disorder, depression, anxiety (generalised anxiety and social anxiety combined), and disruptive behaviour disorders (DBD; encompassing oppositional defiant and conduct disorders). The anxiety and DBD phenotypes were created to reflect the fact that they share diagnostic boundaries. We focussed on these six as they clustered into three latent dimensions, are relatively common among autistic individuals, and because there are well powered external GWAS enabling us to study the impact of PGS on these phenotypes in autistic individuals.

For these six phenotypes, we calculated the prevalence in SPARK and compared them with known prevalence among autistic individuals and in the general population, and assessed their co-occurrence using tetrachoric correlation.

Given the correlation among these phenotypes, we assessed if there are underlying latent structures, further details of which are provided in the **Supplementary Methods** and **Supplementary Figure 3**. Confirmatory factor analysis was carried out with the lavaan package (version 0.6.15) in R³⁴.

2.3. Quality control of genetic data

Genotype data quality control and analyses was restricted to individuals of genetically inferred European ancestries, as provided by the SPARK team. Prior to imputation, we removed variants with less than 95% genotyping rate and those not in Hardy-Weinberg Equilibrium ($p < 1E-6$). We then excluded individuals with genotyping rate less than 98%, and where family data was available, with excess in Mendelian errors. We additionally excluded individuals with excess heterozygosity (greater than 3 standard deviations from the mean), and whose chromosomal sex did not match their reported sex assigned at birth, as these could indicate technical errors. We imputed participants using the TOPMED imputation panel³⁵, using the Michigan Imputation Server³⁶ (Phasing using Eagle v2.4³⁷ and imputation using Minimac4³⁸). After imputation, we

retained SNPs with a minor allele frequency $> 0.5\%$ and an imputation $r^2 > 0.6$. We used liftOver³⁹ to convert from GRCh38/hg38 to GRCh37/hg19 build.

Following imputation, plink-format data including 9,698,250 variants for 47,170 participants (autistic and non-autistic) in the WES1-4 component of iWES v2 was obtained. Of this, data for 17,582 autistic individuals were kept for analysis; these had a total genotyping rate of 0.997. Filters for SNP missingness (92% threshold) identified 56,853 SNPs with high missingness, which were excluded. All individuals were within the 0.2 threshold for inbreeding coefficient. As a final step, 1,174 variants were removed due to Hardy-Weinberg exact test ($p < 1E-6$). This left 9,640,223 SNPs which were used for subsequent analysis.

2.4. Generation of genetic principal components

To calculate genetic principal components, we first pruned for linkage disequilibrium (LD) [plink⁴⁰ --indep-pairphase 50 5 0.2], after which SNPs in regions with long range LD were identified and removed. Individual relatedness was calculated using KING⁴¹ after which principal components were calculated with PC-AIR in GENESIS^{42,43} which accounts for relatedness among individuals.

2.5. SNP heritability and genetic correlations

SNP heritabilities and genetic correlations for factor scores and individual phenotypes were computed using single-component GCTA-GREML⁴⁴ (GREML-SC) and Linkage Disequilibrium Score Regression (LDSC)⁴⁵. GREML analysis used individual-level data with a pruned GRM, while LDSC used summary statistics. Heritability estimates were calculated on both observed and liability scales, with population prevalence estimates obtained from Lai et al., 2019². Genetic correlation among categorical conditions and factors in autistic individuals were calculated using GCTA-GREML. Genetic correlation between the factor scores and diagnostic categories with out-of-sample summary statistics for autism from the non-overlapping iPSYCH dataset¹⁶, anxiety⁴⁶, ADHD²⁶, schizophrenia⁴⁷, bipolar disorder⁴⁸, PTSD⁴⁹, depression⁵⁰, and DBD among individuals with ADHD⁵¹ were calculated using LDSC. Further details are provided in the **Supplementary Methods**.

2.6. GWAS of factor scores and co-occurring mental health conditions

GWAS was performed separately for each factor (using factor scores) and co-occurring mental health condition. This was done using fastGWA-GLMM^{52,53} in the GCTA package version v1.94.1 Linux. FastGWA-GLMM implements a linear mixed effects model for traits in a resource-efficient manner. Each GWAS was performed with a total of 17,573 individuals (from 15,798 families) and for 8,534,720 SNPs with missingness rate < 0.10. Covariates included age at registration (to account for the fact that some mental health conditions are diagnosed only in adolescence or later), sex, and the first 10 genetic principal components. Manhattan and quantile-quantile plots are provided in **Supplementary Figure 4**.

2.7. Polygenic scoring

Polygenic scoring was done using PRSCs⁵⁴ which utilises Bayesian regression framework: inference of posterior effect sizes of SNPs is made by using genome-wide association summary statistics and an LD reference panel which is external to the test dataset. Polygenic scores (PGS) were generated for 11 phenotypes using summary statistics from the cited sources: The iPSYCH-2015 autism GWAS¹⁶, ADHD²⁶, schizophrenia⁴⁷, bipolar disorder⁴⁸, depression⁵⁰, anxiety disorder⁴⁶, neuroticism⁵⁵, P factor³², intelligence⁵⁶ and educational attainment⁵⁷, with hair colour⁵⁸ as a negative control. We restricted PGS analyses to 13,290 genetically unrelated autistic individuals identified using GCTA-GREML.

2.8. Copy number variant and *de novo* variant data

We generated copy number variant (CNV) data from SNP arrays, and filters were applied as in Huguet et al (2021)⁵⁹. All genes fully encompassed in CNVs were annotated with constraint scores. CNV burden scores were calculated as the count of highly constrained (LOEUF topmost decile)⁶⁰ genes, separately for deletions and duplications (see **Supplementary Methods** for details).

De novo variants (DNV) were obtained from trios in SPARK as previously described⁶¹. DNV analysis was limited to 4,976 autistic individuals with available data. For analysis, protein truncating variants (PTVs) were defined as those with one of these coding consequence:

"frameshift_variant", "start_lost", "stop_gained", "splice_donor_variant", "splice_acceptor_variant" or "stop_lost". Missense variants were filtered for those with an MPC score ≥ 2 . Both PTVs and missense variants were further subset to those with Loss-of-function Observed/Expected Upper-bound Fraction (LOEUF) scores in the topmost decile, and individuals were assigned a binary score for their carrier status. Both PTVs and missense variants were combined to create the category of constrained DNVs.

2.9. Regression analyses with PGS, CNV and DNV data

Factor scores were scaled using RobustScaler in ScikitLearn⁶², which scales features using methods that are robust to outliers. Logistic regression analysis within the generalised linear model was performed separately for each phenotype of interest, with PGS, DNVs and CNVs in separate models. In the baseline model, we first tested the association between the phenotypes and age at registration, sex, and cognitive impairment. Next, univariate analyses were conducted to test the association of each genetic variable with each factor and categorical condition separately. This was done over three models: Model 1 with age and the first ten genetic principal components as covariates; Model 2 with sex and sex*genetic variable interaction added (to investigate if the effects differ by sex); and Model 3 with presence/absence of cognitive impairment and its interaction with the genetic variable added (to investigate if the effects differ by co-occurring cognitive impairment). See **Supplementary Methods** for further detail. As a positive control, we also investigated the association between DNVs and CNVs with cognitive impairment.

Thereafter, to assess the combined effect of multiple genetic variables, ten PGS (excluding hair colour) with variance inflation factor not exceeding 4 were included in a multiple PGS model for each factor and categorical condition, with age, sex, cognitive impairment, and the first 10 genetic principal components included as covariates.

In addition, given that for phenotypes such as schizophrenia, bipolar disorder and depression, the younger participants had either not yet reached or were still within the age of risk, Cox proportional hazards regression was also utilised in further sensitivity analyses, with each

categorical condition as outcome variable and sex, cognitive impairment, and the first ten genetic principal components included as covariates (see **Supplementary Methods**).

To explore direct and indirect genetic effects (i.e., genetic effects mediated through parental familial effects or nurture), a subset of trios (n= 5,236) for which PGS were available for probands, mothers and fathers were reanalysed in regression models with each factor and phenotype separately as outcome of interest. Analysis was done separately only for PGS with significant effects in the baseline model, with the corresponding PGS for both parents included in each model as previously done⁶³. We included age, sex, cognitive impairment and the first 10 genetic principal components derived from the child's genotypes as covariates.

Across all analyses, multiple testing correction was done using Benjamini-Yekutieli method of false discovery rate.

2.10. Ethics statement

The Cambridge University Human Biology Research Ethics Committee provided ethical approval to access and analyse de-identified data from SPARK. The SPARK study is funded by the Simons Foundation and uses a single, central Institutional Review Board (WCG IRB Protocol #20151664). Informed consent was obtained by the SPARK study team from the participant or the caregiver/parent.

3. Results

3.1. Phenotypic profiles of co-occurring health conditions in autistic individuals

Among autistic individuals in the SPARK cohort (n=74,204; median age at registration = 9 years [IQR = 9 years], 26% female, 56% having a co-occurring psychiatric condition and 16.3% with intellectual disability or cognitive impairment), the prevalence of the six mental health conditions was higher than the general population, but similar to the prevalence previously published in autistic populations (Lai et al, 2019) (**Supplementary Table 1**). After accounting for the age of the participants, these conditions exhibited significant sex differences - with anxiety,

depression and bipolar disorder more prevalent among females and ADHD more prevalent among males. With the exception of schizophrenia, all conditions were significantly more prevalent among autistic individuals without co-occurring cognitive impairment (**Supplementary Table 1**). We subsequently investigated the correlation and co-occurrence among the six mental health conditions. We identified correlations ranging from 0.63 (ADHD and DBD, and schizophrenia and bipolar disorder) to 0.19 (ADHD and schizophrenia) (**Figures 1a and b**).

Given this correlation, we conducted confirmatory factor analyses to identify latent factors. The best performing model (**Supplementary Table 2** and **Supplementary Figure 2**) included three factors - a behavioural factor (including ADHD and disruptive behaviour disorders), a cothymic factor (comprising anxiety and depression), and a 'Kraepelin' or thought disorder factor which grouped schizophrenia and bipolar disorder together (CFI = 0.972, TLI = 0.929, RMSEA = 0.05).

3.2. Common variant genetic architecture of co-occurring mental health conditions

To understand if common genetic variants are associated with co-occurring mental health conditions and factors, we calculated SNP heritabilities on the liability scale using GREML-SC and LDSC (**Supplementary Table 3** and **Figure 1C and 1D**). Both methods identified moderate and significant SNP heritabilities for behavioural conditions (ADHD and DBD) and the behavioural factor. In addition, we identified significant but low SNP heritabilities for cothymic and thought disorder factors. The heritability of the other phenotypes failed to reach statistical significance, potentially due to the low statistical power and underdiagnosis of younger autistic individuals.

We further investigated the genetic correlation among these phenotypes. We identified high genetic correlations among the 3 factors (**Supplementary Table 4**). Using GREML-SC, among categorical conditions, we identified high and significant genetic correlation among DBD, anxiety, and ADHD, and high genetic correlation between bipolar disorder and both DBD and anxiety. Estimates of these genetic correlations were consistent between GREML-SC and LDSC.

Subsequently, we calculated genetic correlations between the co-occurring factors and conditions, and seven external GWAS of psychiatric conditions, including autism. The external

GWAS for ADHD, DBD, PTSD, and depression had significant and moderate genetic correlations with the behavioural factor, cothymic factor, ADHD and DBD in autistic individuals (**Supplementary Table 5 and Figure 2**).

3.3. Genetic associations of factor scores and co-occurring mental health conditions

To further investigate the link between genetic variants and co-occurring mental health phenotypes, we ran regression between genetic variants (CNV scores, DNVs, and PGS) and the mental health phenotypes. In univariate models (Model 1), PGS for ADHD, depression, and educational attainment were significantly associated with all phenotypes in the expected direction (**Supplementary Tables 6 and 8, Figure 3**). We obtained consistent results when accounting for cognitive impairment (Model 3), and consistent results for the three factors, DBD, and ADHD when accounting for sex (Model 2). In addition, we observed significant associations between PGS for anxiety and IQ for several of these phenotypes. Our control PGS (hair colour) was not associated with any phenotypes.

While rare genetic variants significantly impacted the likelihood of cognitive impairment as expected (**Figure 4 and Supplementary Figure 5**), with the largest effects observed for *de novo* PTV variants followed by deletions, *de novo* missense variants, and duplications, we identified no significant association between any of these rare genetic variants and any of the mental health phenotypes. We obtained largely consistent results when using Cox proportional hazard regression (**Supplementary Table 9**), and identified no significant interaction with sex or cognitive impairment (**Supplementary Tables 6 and 8**).

We followed up the univariate analyses with multivariate analyses to account for the correlation among polygenic scores. In these, we simultaneously regressed ten PGS against the outcome phenotypes. In these conditional analyses, PGS for ADHD, depression, and EA were associated in the expected direction with all factors (**Supplementary Table 7**) and the majority of the mental health conditions (**Supplementary Table 10**). Yet again, we obtained consistent results when using Cox proportional-hazards regression.

We further wondered if the association between PGS for EA, ADHD and depression represent direct genetic effects of the child, or indirect genetic effects of parental nurture. We tested this using models of PGS association including proband and parent PGS for each phenotype (**Supplementary Tables 11 and 12**, and **Figure 5**). We found no evidence of indirect effects - there was no significant association between either paternal or maternal PGS and any of the phenotypes tested. However, we found significant association between the autistic proband's depression and EA PGS and cothymic and behavioural factors even after conditioning on parental PGS. In addition, the autistic proband's ADHD PGS was significantly associated with the behavioural factor, and EA PGS with DBD. All associations were in the expected direction.

4. Discussion

In this study, we conducted the first large-scale genetic investigation of co-occurring mental health phenotypes among autistic individuals from the SPARK cohort. The prevalence of the studied mental health conditions in autistic individuals approximated what has been reported in literature², albeit with deviations for schizophrenia (lower prevalence) and ADHD (higher prevalence). This may reflect the demographic profile of participants in the SPARK cohort.

We highlight five key results. First, we demonstrate that ADHD, DBD, and all mental health factors have modest and significant SNP heritabilities, suggesting that these emerge from complex sociobiological mechanisms. Although other conditions tested did not have significant SNP heritabilities, this may be because of low statistical power (e.g., schizophrenia), and because the majority of autistic individuals are young and may not have reached an appropriate age for diagnosis. Second, we identify moderate to high genetic correlations between co-occurring ADHD, DBD, anxiety and all three mental health factors among autistic individuals and the corresponding conditions in external population cohorts. This suggests similar genetic variants are associated with mental health phenotypes both among autistic individuals and external cohorts. Third, neither PGS analyses nor genetic correlation studies identify shared genetics between autism and co-occurring mental health phenotypes specifically among autistic

individuals. There is a positive and moderate genetic correlation between autism and other mental health conditions in the external population cohorts, suggesting that increased polygenic propensity for autism is associated with higher likelihood of mental health conditions. Taken together, this suggests that co-occurring mental health conditions in autistic individuals arise from a combination of their elevated autism polygenic propensity and additional mental health-specific genetic factors in addition to other socio-contextual factors. Fourth, within family analyses indicate that the effects of the PGS on some mental health phenotypes are direct (i.e., effects of inherited genetic variation on the child's phenotype) and we find no evidence for indirect genetic effects (i.e., parental genetics influencing the autistic child's mental health through the familial environment). Fifth, and finally, we find no evidence for the role of rare genetic variants on increasing the likelihood for co-occurring mental health conditions among autistic individuals, despite an association between rare variants and mental health conditions in external population cohorts^{28,64-66}. This pattern may reflect collider bias, limited statistical power, phenotypic heterogeneity between autistic individuals with and without rare variants¹⁶, or diagnostic overshadowing of mental health conditions^{67,68} in autistic individuals with rare variants who may have complex support needs. The finding of heritable latent factors suggests the utility of a biaxial system of diagnosis, with dimensional latent factors on one axis and conventional categorical diagnosis on another.

For oppositional defiant and conduct disorder, which together form the kernel of disruptive behaviour disorders in ICD-11⁶⁹, there was no comparable SNP heritability in external population cohorts. Notably, the GREML-SC and LDSC based SNP heritability for DBD was high: 0.52 (se: 0.08) in line with a previous study on the SNP heritability for DBD among individuals with ADHD ($h^2_{\text{SNP}} = 0.34$; $SE = 0.05$)⁵¹. The finding of significant genetic correlation between disruptive behaviour disorders in the SPARK and iPSYCH cohorts suggests the two constructs are consistent. As observed by Sesso et al (2020)⁷⁰, autism and disruptive behaviour disorders represent distinct phenotypes among individuals with ADHD. Given the high heritability of DBD among autistic individuals, it may be possible that autistic individuals with co-occurring DBD represent a distinct subgroup.

PGS for ADHD and depression were significantly associated with several phenotypes. However, among all the PGS, the most consistent association with all the co-occurring phenotypes was with educational attainment, reflecting the shared genetics between educational attainment and a wide number of mental health conditions in population cohorts⁷¹.

This study has a few limitations. First, it was conducted with only one dataset, which may limit its generalisability. Future work will seek to replicate the study in other comparable datasets. Second, there was also a modest number of participants with some conditions - especially schizophrenia and bipolar disorder - which may have accounted for the lack of precision of some estimates. The low sample size also meant that we had limited statistical power to detect more modest effects for DNVs and in trio-based analyses of PGS. Furthermore, mental health conditions may be underdiagnosed in carriers of DNVs who often have complex support needs. Third, other factors which may have influenced the heritability and genetic correlation estimates include sample overlap, assortative mating and collider bias. Specifically, for depression, we were unable to estimate genetic correlation using LDSC both among autistic individuals and with the external population due to the low SNP heritability. Fourth, we were able to investigate the genetic correlates of mental health conditions only among individuals of genetically inferred European ancestries, and the generalisability of these findings to other population groups is unknown. Furthermore, for depression, we were unable to estimate genetic correlation using LDSC both among autistic individuals and with the external population due to the low SNP heritability.

ACKNOWLEDGEMENTS

This study was conducted as part of AO's doctoral programme at the University of Cambridge, funded by the Cambridge Trust and a Trinity-Henry Barlow studentship. SBC received funding from the Wellcome Trust 214322\Z\18\Z. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. SBC also received funding from the Innovative Medicines Initiative 2 Joint

Undertaking under grant agreement No 777394 for the project AIMS-2-TRIALS. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. SBC also received funding from the Autism Centre of Excellence (ACE) at Cambridge, SFARI, the Templeton World Charitable Fund, the MRC, and the NIHR Cambridge Biomedical Research Centre. The research was supported by the National Institute for Health Research (NIHR) Applied Research Collaboration East of England. Any views expressed are those of the author(s) and not necessarily those of the funder. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. Any views expressed are those of the author(s) and not necessarily those of the funders (including IHI-JU2, the NIHR and the Department of Health and Social Care).

References

1. Lai MC. Mental health challenges faced by autistic people. *Nature human behaviour*. 2023;7(10). doi:10.1038/s41562-023-01718-2
2. Lai MC, Kassee C, Besney R, et al. Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The lancet Psychiatry*. 2019;6(10). doi:10.1016/S2215-0366(19)30289-5
3. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry*. 2008;47(8):921-929.
4. Croen LA, Zerbo O, Qian Y, et al. The health status of adults on the autism spectrum. *Autism*. 2015;19(7):814-823.
5. Lever AG, Geurts HM. Psychiatric Co-occurring Symptoms and Disorders in Young, Middle-Aged, and Older Adults with Autism Spectrum Disorder. *J Autism Dev Disord*. 2016;46(6):1916-1930.
6. Doherty M, Neilson S, O'Sullivan J, et al. Barriers to healthcare and self-reported adverse outcomes for autistic adults: a cross-sectional study. *BMJ Open*. 2022;12(2):e056904.
7. Liu BM, Paskov K, Kent J, et al. Racial and Ethnic Disparities in Geographic Access to Autism Resources Across the US. *JAMA Netw Open*. 2023;6(1):e2251182.
8. Wallisch A, Boyd BA, Hall JP, et al. Health Care Disparities Among Autistic LGBTQ+ People. *Autism Adulthood*. 2023;5(2):165-174.
9. Department for Work. Landmark review calls on employers to boost support for autistic people. GOV.UK. February 28, 2024. Accessed November 7, 2024. <https://www.gov.uk/government/news/landmark-review-calls-on-employers-to-boost-support-for-autistic-people>
10. Boilson AM, Churchard A, Connolly M, Casey B, Sweeney MR. Screening for Autism Spectrum Condition Through Inner City Homeless Services in the Republic of Ireland. *Journal of Autism and Developmental Disorders*. 2022;53(10):3987-3998.
11. Churchard A, Ryder M, Greenhill A, Mandy W. The prevalence of autistic traits in a homeless population. *Autism*. Published online April 10, 2018. doi:10.1177/1362361318768484
12. Griffiths S, Allison C, Kenny R, Holt R, Smith P, Baron-Cohen S. The Vulnerability Experiences Quotient (VEQ): A Study of Vulnerability, Mental Health and Life Satisfaction in Autistic Adults. *Autism Res*. 2019;12(10):1516-1528.

13. Stack A, Lucyshyn J. Autism Spectrum Disorder and the Experience of Traumatic Events: Review of the Current Literature to Inform Modifications to a Treatment Model for Children with Autism. *J Autism Dev Disord*. 2019;49(4):1613-1625.
14. Grove J, Ripke S, Als TD, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019;51(3):431-444.
15. Tick B, Bolton PF, Happé F, Rutter M, Rijdsdijk F. Heritability of autism spectrum disorders: A meta-analysis of twin studies. *J Child Psychol Psychiatry*. 2016;57(5):585-595.
16. Warriar V, Zhang X, Reed P, et al. Genetic correlates of phenotypic heterogeneity in autism. *Nat Genet*. 2022;54(9):1293-1304.
17. Havdahl A, Niarchou M, Starnawska A, Uddin M, van der Merwe C, Warriar V. Genetic contributions to autism spectrum disorder. *Psychol Med*. 2021;51(13):2260-2273.
18. Antaki D, Guevara J, Maihofer AX, et al. Publisher Correction: A phenotypic spectrum of autism is attributable to the combined effects of rare variants, polygenic risk and sex. *Nat Genet*. 2022;54(8):1259.
19. LaBianca S, LaBianca J, Pagsberg AK, et al. Copy Number Variants and Polygenic Risk Scores Predict Need of Care in Autism and/or ADHD Families. *Journal of Autism and Developmental Disorders*. 2020;51(1):276-285.
20. Yap CX, Alvares GA, Henders AK, et al. Analysis of common genetic variation and rare CNVs in the Australian Autism Biobank. *Molecular Autism*. 2021;12(1):1-17.
21. Cirnigliaro M, Chang TS, Arteaga SA, et al. The contributions of rare inherited and polygenic risk to ASD in multiplex families. *Proceedings of the National Academy of Sciences*. 2023;120(31):e2215632120.
22. Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS. Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry*. 2018;24(3):409-420.
23. Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45(9):984-994.
24. Polderman TJC, Benyamin B, de Leeuw CA, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*. 2015;47(7):702-709.
25. Warriar V, Baron-Cohen S. Childhood trauma, life-time self-harm, and suicidal behaviour and ideation are associated with polygenic scores for autism. *Mol Psychiatry*. 2021;26(5):1670-1684.
26. Demontis D, Walters GB, Athanasiadis G, et al. Genome-wide analyses of ADHD identify 27

- risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet.* 2023;55(2):198-208.
27. Zhang X, Grove J, Gu Y, et al. An axis of genetic heterogeneity in autism is indexed by age at diagnosis and is associated with varying developmental and mental health profiles. *medRxiv*. Published online 2024:2024-2007.
 28. Satterstrom FK, Walters RK, Singh T, et al. Autism spectrum disorder and attention deficit hyperactivity disorder have a similar burden of rare protein-truncating variants. *Nat Neurosci.* 2019;22(12):1961-1965.
 29. Rees E, Creeth HDJ, Hwu HG, et al. Schizophrenia, autism spectrum disorders and developmental disorders share specific disruptive coding mutations. *Nature Communications.* 2021;12(1):1-9.
 30. Chick SL, Holmans P, Grozeva D, et al. Analysis of exome sequencing data implicates rare coding variants in STAG1 and ZNF136 in schizophrenia. *medRxiv*. Published online October 16, 2024:2024.10.14.24314843. doi:10.1101/2024.10.14.24314843
 31. Plana-Ripoll O, Pedersen CB, Holtz Y, et al. Exploring Comorbidity Within Mental Disorders Among a Danish National Population. *JAMA Psychiatry.* 2019;76(3):259-270.
 32. Grotzinger AD, Rhemtulla M, de Vlaming R, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav.* 2019;3(5):513-525.
 33. SPARK Consortium. Electronic address: pfeliciano@simonsfoundation.org, SPARK Consortium. SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research. *Neuron.* 2018;97(3):488-493.
 34. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *J Stat Soft.* 2012;48:1-36.
 35. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature.* 2021;590(7845):290-299.
 36. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nature Genetics.* 2016;48(10):1284-1287.
 37. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nature Genetics.* 2016;48(11):1443-1448.
 38. Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics.* 2015;31(5):782-784.
 39. Hinrichs AS, Karolchik D, Baertsch R, et al. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* 2006;34(Database issue):D590-D598.

40. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559-575.
41. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics.* 2010;26(22):2867-2873.
42. Gogarten SM, Sofer T, Chen H, et al. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics.* 2019;35(24):5346-5348.
43. Conomos MP, Thornton T. GENetic ESTimation and inference in structured samples (GENESIS): statistical methods for analyzing genetic data from samples with population structure and/or relatedness. *R package version.* 2016;2(0.1).
https://www.imsbio.co.jp/RGM/R_rdfile?f=GENESIS/man/GENESIS-package.Rd&d=R_BC
44. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010;42(7):565-569.
45. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics.* 2015;47(3).
doi:10.1038/ng.3211
46. Purves KL, Coleman JRI, Meier SM, et al. A major role for common genetic variation in anxiety disorders. *Mol Psychiatry.* 2020;25(12):3292-3303.
47. Trubetsky V, Pardiñas AF, Qi T, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature.* 2022;604(7906):502-508.
48. Mullins N, Forstner AJ, O'Connell KS, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet.* 2021;53(6):817-829.
49. Nievergelt CM, Maihofer AX, Atkinson EG, et al. Genome-wide association analyses identify 95 risk loci and provide insights into the neurobiology of post-traumatic stress disorder. *Nat Genet.* 2024;56(5):792-808.
50. Als TD, Kurki MI, Grove J, et al. Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat Med.* 2023;29(7):1832-1844.
51. Demontis D, Walters RK, Rajagopal VM, et al. Risk variants and polygenic architecture of disruptive behavior disorders in the context of attention-deficit/hyperactivity disorder. *Nature Communications.* 2021;12(1):1-12.
52. Jiang L, Zheng Z, Fang H, Yang J. A generalized linear mixed model association tool for biobank-scale data. *Nat Genet.* 2021;53(11):1616-1621.

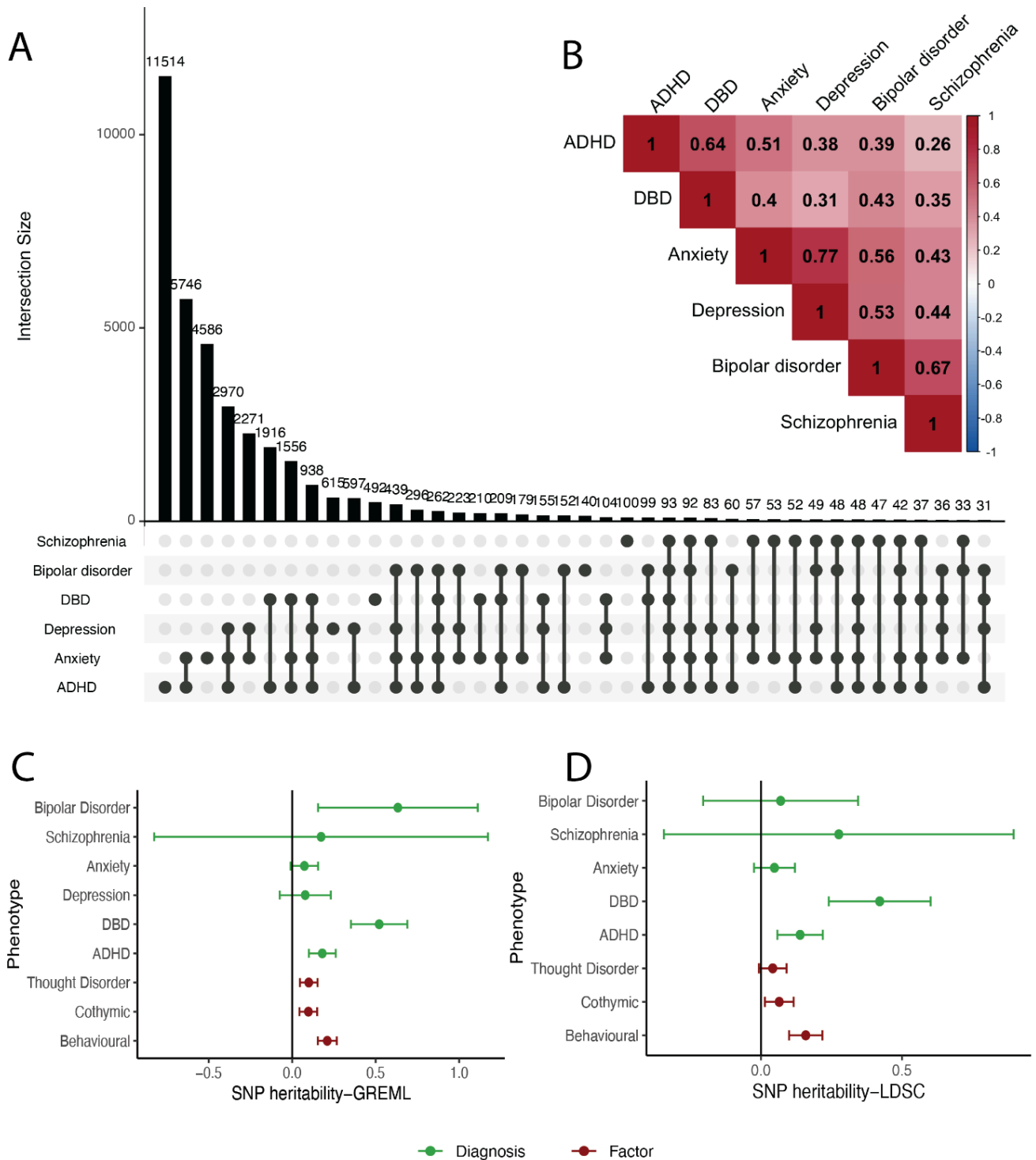
53. Jiang L, Zheng Z, Qi T, et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat Genet.* 2019;51(12):1749-1755.
54. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun.* 2019;10(1):1776.
55. Nagel M, Jansen PR, Stringer S, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet.* 2018;50(7):920-927.
56. Savage JE, Jansen PR, Stringer S, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet.* Published online June 25, 2018:1.
57. Okbay A, Wu Y, Wang N, et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet.* 2022;54(4):437-449.
58. Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet.* 2019;51(9):1339-1348.
59. Huguet G, Schramm C, Douard E, et al. Genome-wide analysis of gene dosage in 24,092 individuals estimates that 10,000 genes modulate cognitive ability. *Molecular psychiatry.* 2021;26(6). doi:10.1038/s41380-020-00985-z
60. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434-443.
61. Koko M, Kyle Satterstrom F, Autism Sequencing Consortium, APEX consortium, Warriar V, Martin H. Contribution of autosomal rare and de novo variants to sex differences in autism. *medRxiv.* Published online April 16, 2024:2024.04.13.24305713. doi:10.1101/2024.04.13.24305713
62. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research.* 2011;12(85):2825-2830.
63. Huang QQ, Wigdor EM, Campbell P, et al. Dissecting the contribution of common variants to risk of rare neurodevelopmental conditions. *medRxiv.* Published online March 6, 2024:2024.03.05.24303772. doi:10.1101/2024.03.05.24303772
64. Palmer DS, Howrigan DP, Chapman SB, et al. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. *Nat Genet.* 2022;54(5):541-547.
65. Singh T, Poterba T, Curtis D, et al. Rare coding variants in ten genes confer substantial risk

for schizophrenia. *Nature*. 2022;604(7906):509-516.

66. Tian R, Ge T, Kweon H, et al. Whole-exome sequencing in UK Biobank reveals rare genetic architecture for depression. *Nature Communications*. 2024;15(1):1-12.
67. The importance of understanding the behavioural phenotypes of genetic syndromes associated with intellectual disability. *Paediatrics and Child Health*. 2014;24(10):468-472.
68. Hagopian LP, Jennett HK. Behavioral Assessment and Treatment of Anxiety in Individuals with Intellectual Disabilities and Autism. *Journal of Developmental and Physical Disabilities*. 2008;20(5):467-483.
69. ICD-11. Accessed November 10, 2024. <https://icd.who.int/en>
70. Sesso G, Cristofani C, Berloffia S, et al. Autism Spectrum Disorder and Disruptive Behavior Disorders Comorbidities Delineate Clinical Phenotypes in Attention-Deficit Hyperactivity Disorder: Novel Insights from the Assessment of Psychopathological and Neuropsychological Profiles. *Journal of Clinical Medicine*. 2020;9(12):3839.
71. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nature Genetics*. 2015;47(11):1236-1241.

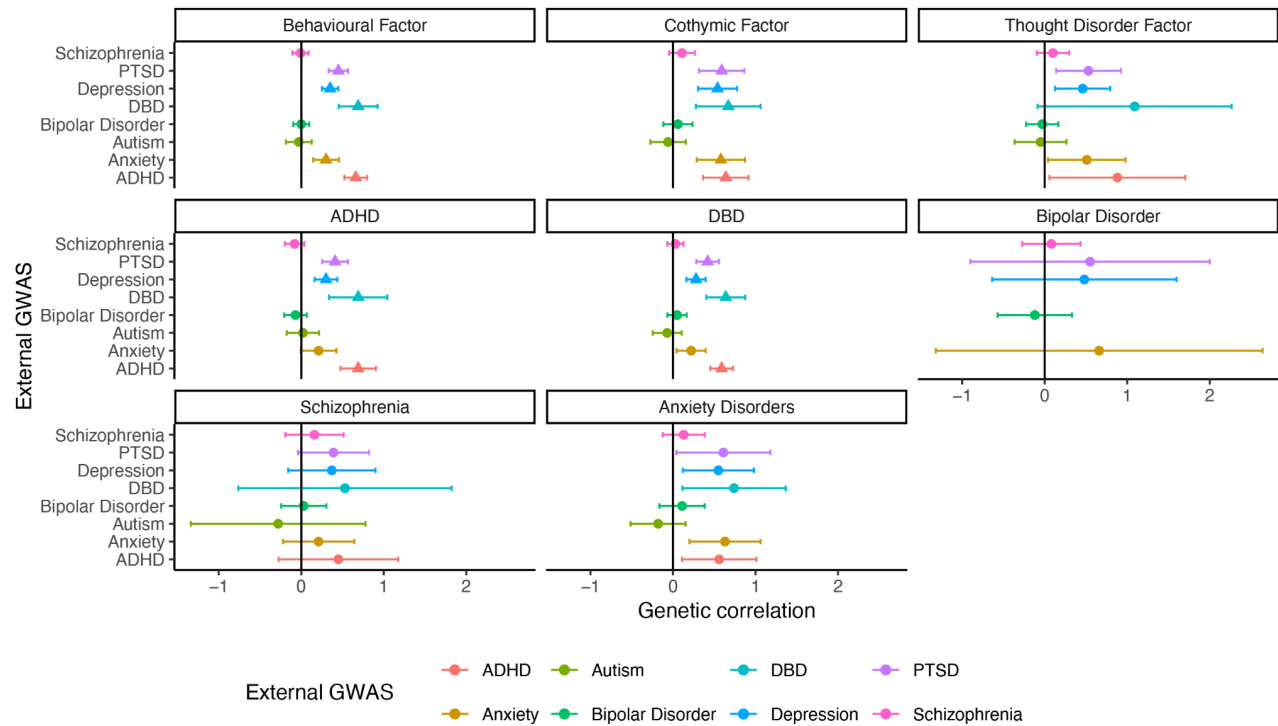
Figure Legends

Figure 1. Phenotypic overlap and correlation among the six conditions and SNP-based heritability.



A: Upset plot showing overlapping diagnoses among autistic individuals. DBD: disruptive behaviour disorder; ADHD: attention deficit hyperactivity disorder. Sample sizes are: ADHD = 27,573, Anxiety = 20,757, Bipolar disorder = 2,633, DBD = 6,375, Depression = 9,231, Schizophrenia = 988. B: Phenotype correlations of the six co-occurring diagnoses. C: SNP heritability estimates (GCTA-GREML) of the diagnoses and factors. Point estimates indicate heritability, and whiskers indicate 95% confidence intervals. D: SNP heritability estimates (LDSC) of the diagnoses and factors. Point estimates indicate heritability, and whiskers indicate 95% confidence intervals.

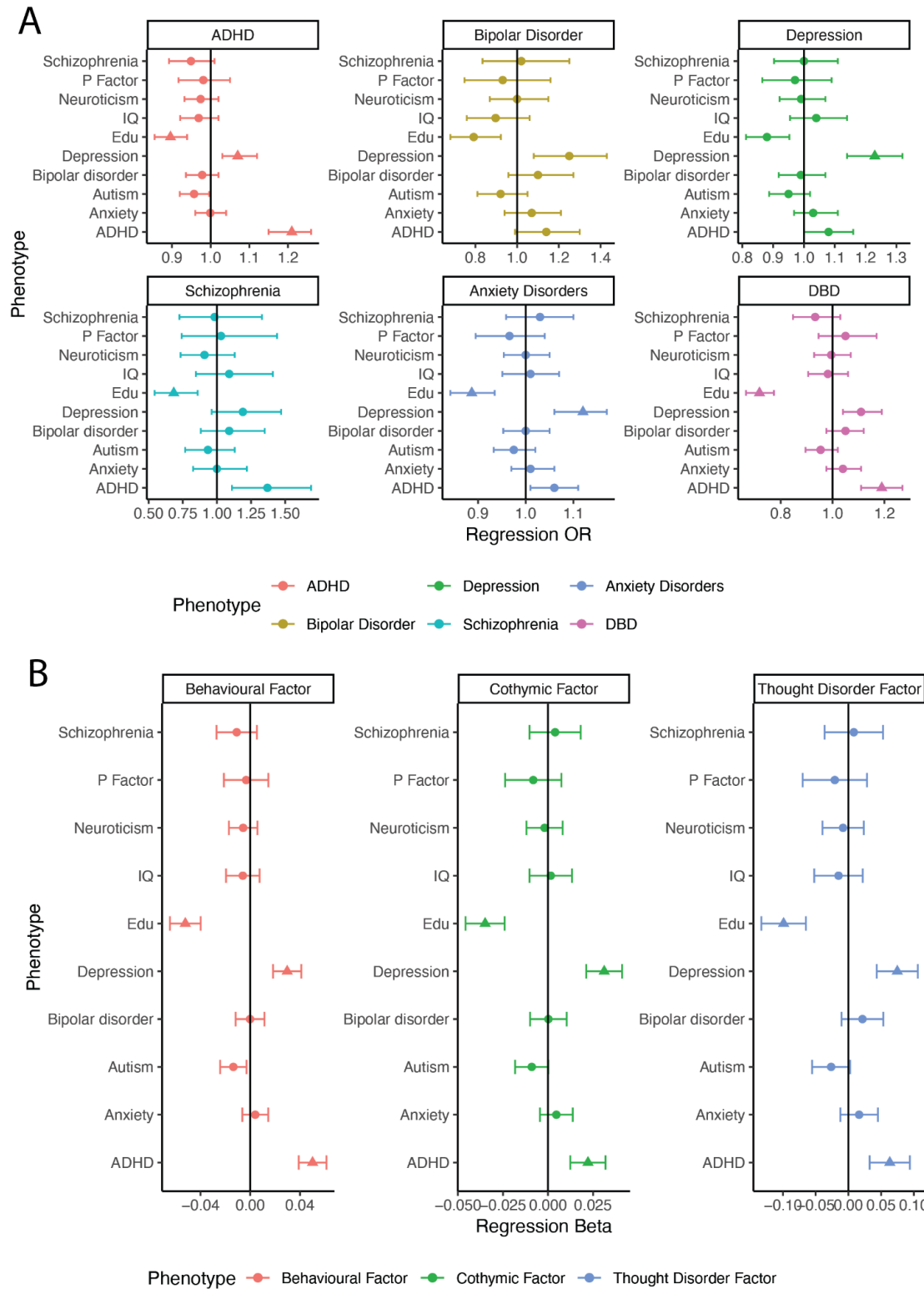
Figure 2. Genetic correlations between mental health phenotypes in autistic individuals and external population cohorts (LDSC).



Point estimates indicate genetic correlation estimates, and whiskers indicate 95% confidence intervals. Triangles indicate significant genetic correlations after Bejamini-Yekutieli (BY) multiple testing correction.

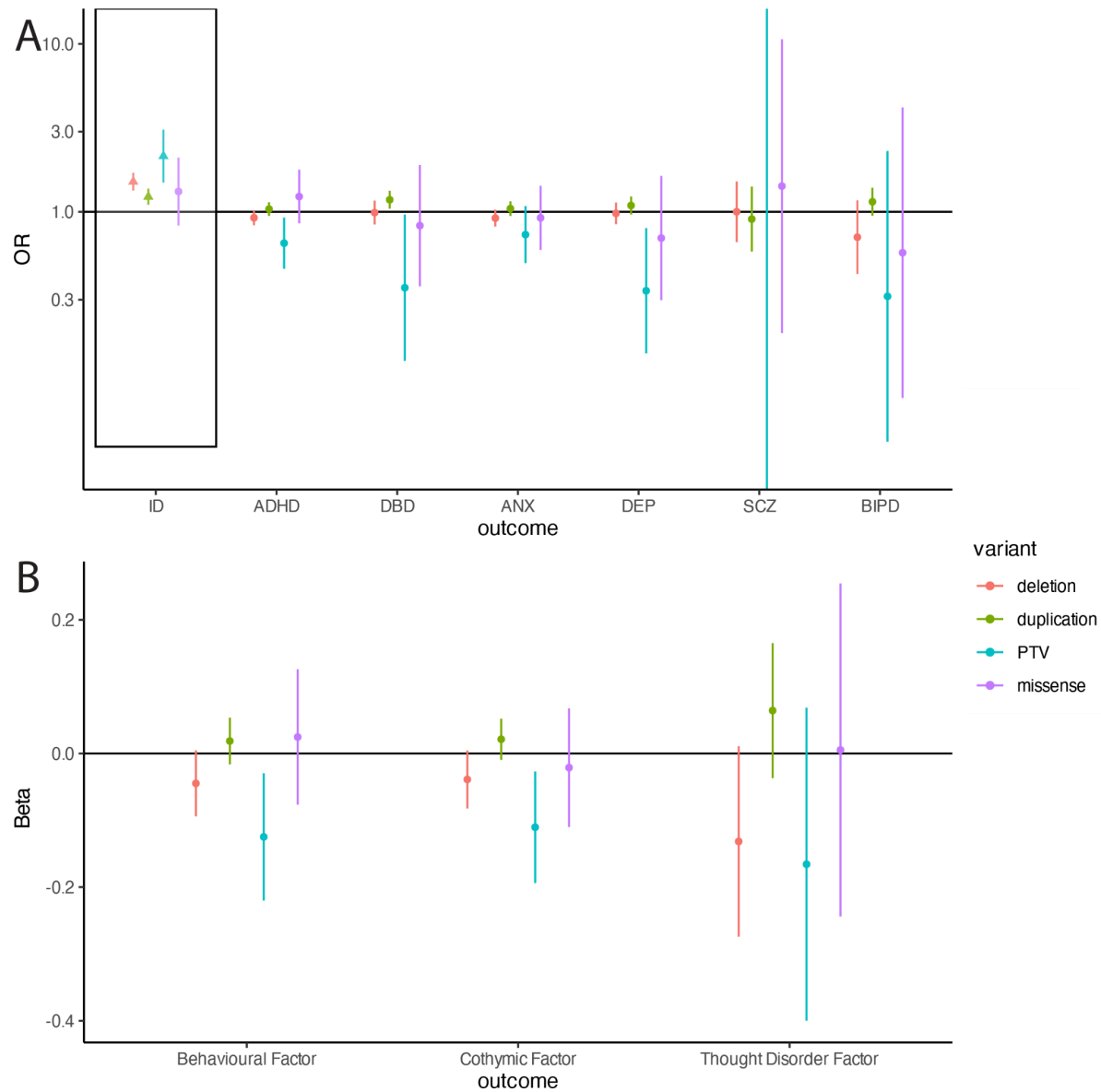
It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Figure 3. Association between PGS and mental health phenotypes among autistic individuals.



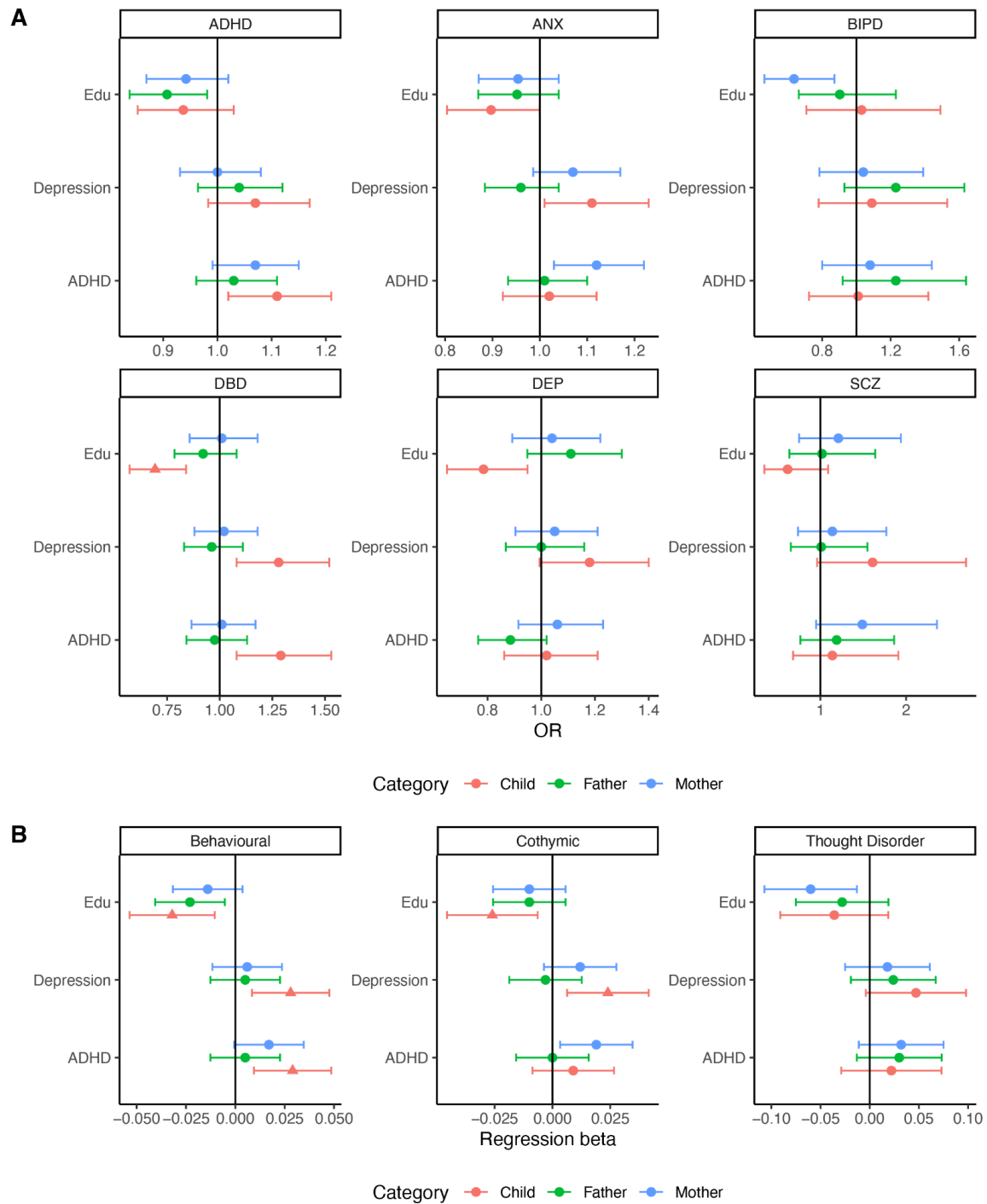
A: Regression with phenotypic categorical diagnoses. DBD: disruptive behaviour disorder; ADHD: attention deficit hyperactivity disorder; DEP: depression; BIPD: bipolar disorder; ANX: anxiety disorder; SCZ: schizophrenia. B: Regression models with robustly scaled factor scores and 10 polygenic scores (PGS). Edu: educational attainment PGS. For both plots, point estimates indicate regression coefficients and whiskers indicate 95% confidence intervals. Triangles represent PGS with significant association after Bejamini-Yekutieli (BY) multiple testing correction.

Figure 4: Association between rare-variants and mental health phenotypes among autistic individuals.



A: Regression models for intellectual disability and categorical diagnoses. B: Regression models for robustly scaled factor scores. ID: intellectual disability. P_{adj} : adjusted p values; ns: not significant. PTV: protein truncating variants. For both plots, point estimates indicate regression coefficients and whiskers indicate 95% confidence intervals. Triangles represent PGS with significant association after Bejamini-Yekutieli (BY) multiple testing correction. ID = intellectual disability, ANX = anxiety, DEP = depression, SCZ = schizophrenia, BIPD = Bipolar disorder

Figure 5: Family (trios) models with PGS.



A: Regression models with categorical diagnoses. B: Regression models with robustly scaled factor scores. For both plots, point estimates indicate regression coefficients and whiskers indicate 95% confidence intervals. Triangles represent PGS with significant association after

Bejamini-Yekutieli (BY) multiple testing correction. PGS from external GWAS provided on the y-axis. ANX = anxiety, DEP = depression, SCZ = schizophrenia, BIPD = Bipolar disorder