Title: Shared polygenetic variation between ASD and ADHD exerts opposite association patterns with educational attainment

Short title: Shared genetic variation between ASD, ADHD and educational attainment

Key words: ADHD, ASD, EA, polygenic overlap

Ellen Verhoef[1,2], Jakob Grove[3-6], Chin Yang Shapland[7-8], Ditte Demontis[3-5], Stephen Burgess[9-10], Dheeraj Rai[11-13], Anders D. Børglum[3-5], Beate St Pourcain[1,11,14]

1. Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
2. International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands
3. The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark
4. Department of Biomedicine (Human Genetics) and Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark
5. Center for Genomics and Personalized Medicine, Aarhus, Denmark
6. Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark
7. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
8. Population Health Sciences, University of Bristol, Bristol, UK
9. MRC Biostatistics Unit, University of Cambridge, Cambridge, UK
10. Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
11. Centre for Academic Mental Health, Bristol Medical School, University of Bristol, Bristol, UK
12. NIHR Biomedical Research Centre, University of Bristol, Bristol, UK
13. Avon and Wiltshire Partnership NHS Mental Health Trust, Bristol, UK

14. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Corresponding authors: Ellen Verhoef and Beate St Pourcain

Postal address: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands

e-mail: Ellen.Verhoef@mpi.nl / Phone: +31 24 3521946

e-mail: Beate.StPourcain@mpi.nl / Phone: +31 24 3521964 / Fax: +31 24 3521213

Requests for reprints should be sent to the corresponding author.

Abstract

Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) are genetically complex neurodevelopmental disorders that often co-occur. Besides some shared genetic aetiology, both conditions display apparent differences in their genetic architectures, especially in their genetic overlap with educational attainment (EA). Here, we investigate ASD-specific, ADHD-specific and cross-disorder genetic associations with EA using a multivariable regression (MVR) framework and genome-wide association summary statistics from large consortia. Our findings show that EA-related polygenic variation is shared across ASD and ADHD, irrespective of opposite association profiles. For example, ASD risk-increasing alleles selected at threshold P_{thr}<0.0015 (N_{SNPs}=1973), showed a 0.009 increase (SE=0.003) in years of education per log-odds ASD liability. Conditional on these effects, the same alleles captured a 0.029 decrease (SE=0.004) in years of education per log-odds ADHD liability. Likewise, ADHD risk-increasing alleles (P_{thr}<0.0015, N_{SNPs}=2717), showed an inverse ADHD-specific association with EA (β=-0.012(SE=0.003)), conditional on positive cross-disorder effects shared with ASD (β=0.022(SE=0.003)). MVR effect sizes increased when variants were restricted to markers tagging both ASD and ADHD, which at P_{thr}<0.0015 (N_{SNPs}=83) implicated many regulatory loci. The discovered inverse ADHD/ASD cross-disorder effects on EA were stronger and larger than for other combinations of psychiatric disorders, but are unlikely to be limited to ASD and ADHD. Thus, shared polygenic ASD/ADHD variation can, involving different combinations of the same risk-increasing alleles, reveal an opposite association profile with EA. These patterns are consistent with inverse EA-related genetic correlations between ASD and ADHD, as supported by simulations, and may affect the detectable net genetic overlap in cross-disorder investigations.

Introduction

Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) are genetically complex childhood-onset neurodevelopmental disorders[1,2] that often co-occur[3]. Approximately 15–25% of individuals with ADHD show ASD symptoms, and about 40–70% of individuals with ASD have a comorbid ADHD symptomatology[3].

Both rare and common genetic variation contribute to ASD and ADHD liability[4–7]. There is increasing evidence from twin and molecular studies[8,9] that suggests genetic links between ASD and ADHD symptoms, both throughout population variation[10–16] and at the extremes[10,17]. The existence of genetic cross-disorder links is further strengthened by the familial co-aggregation of both clinical disorders in large register-based studies[18]. Consistently, the identification of shared copy number variations (CNVs) in ASD and ADHD suggests similar biological pathways[19]. Estimates of cross-disorder genetic correlations range between 0.54 and 0.87 in twin analyses[20]. For common genetic variation, cross-trait genetic correlations reach up to one in population-based samples[11], but are surprisingly low between clinically defined ASD and ADHD[21–23]. While recent research has reported moderate genetic overlap between clinical ASD and ADHD using genome-wide analyses (r_g=0.36)[21], earlier, less powerful, studies found little evidence for genetic correlation[22,23].

Besides this shared genetic aetiology, one of the most striking differences in the polygenic architecture of clinical ASD and clinical ADHD is the opposite genetic correlation of each disorder with cognitive functioning and educational attainment (EA), a proxy of cognitive abilities and socio-economic status (SES)[24]. While increased polygenic ADHD risk has been linked with lower cognitive abilities and EA[25–29], increased polygenic ASD risk has been associated with higher cognitive functionality and SES[21,26,28,30,31]. This opposite pattern of association is most notable for measures of years of schooling and college completion[21,29]. It is known that ADHD is associated with multiple indicators of socio-economic disadvantage including poverty, housing tenure, maternal education, income, lone parenthood and younger motherhood[32], while ASD, especially among children in the United States, can be positively

associated with socio-economic position[33]. However, opposite association patterns with EA may also reflect ASD-specific and ADHD-specific cognitive and neurological features[34].

Here, we aim to quantify the genetic overlap between ASD, ADHD, and EA. We dissect polygenic associations between each disorder and EA into either ASD-specific or ADHD-specific associations conditional on genetic influences shared across both disorders and EA. The presence of such cross-disorder genetic associations would suggest indistinguishable evidence for pleiotropic, mediating or confounding factors between genetically predictable ASD, ADHD and EA. Their absence would be consistent with the independence of EA-related polygenic variation within ASD and ADHD genetic architectures.

Using a multivariable regression (MVR) framework, we decouple the selection of disorder-related genetic variants from observed genetic effects in ASD, ADHD and EA using genome-wide association study (GWAS) summary statistics from the Danish Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), the Psychiatric Genomics Consortium (PGC), the Social Science Genetic Association Consortium (SSGAC) and the Complex Trait Genetics (CTG) lab. We integrate our findings with current knowledge on genetic interrelationships between ASD, ADHD and EA and propose a multi-factorial model supported by simulations.

Methods and Materials

Genome-wide association study summary statistics

EA and general intelligence: GWAS summary statistics for EA as years of schooling (discovery and replication sample combined, excluding 23andMe)[35] were obtained from the SSGAC (https://www.thessgac.org/, Table 1). GWAS summary statistics of a meta-analysis on general intelligence[28] were retrieved from the CTG lab (https://ctg.cncr.nl/software/summary_statistics, Table S1).

ASD: GWAS summary statistics were accessed through iPSYCH (http://ipsych.au.dk/) using samples from the Danish Neonatal Screening Biobank hosted by

Statens Serum Institute[21,36]. For MVR analyses, we studied an ASD sample excluding patients with a comorbid ADHD diagnosis (ASD without ADHD (ASD(iPSYCH,woADHD), Table 1) to avoid any case overlap with the ADHD GWAS analysis (see below). For genetic correlation analyses, we also studied summary statistics of the full ASD(iPSYCH) sample (Table 1). Independent ASD summary statistics were obtained from the PGC (ASD(PGC), www.med.unc.edu/pgc/, Table 1)[37].

ADHD: Association results for ADHD were obtained through the iPSYCH project[29,36] (ADHD, Table 1). Note that ADHD cases may also have an additional diagnosis of ASD, but do not overlap with cases in ASD(iPSYCH,woADHD)(see above). Controls are, to a large extent, shared among iPSYCH samples.

ASD+ADHD: For genetic correlation analyses, ASD+ADHD summary statistics were derived by conducting a random-effect meta-analysis of ASD(PGC) and ADHD(iPSYCH) summary statistics using GWAMA software[38]. Note that the sample size for ADHD(iPSYCH) is about three times larger than for ASD(PGC) (Table1).

Other psychiatric disorders: GWAS summary statistics for Major Depressive Disorder (MDD)[39], Schizophrenia (SCZ)[40] and Bipolar Disorder (BD)[40] were obtained from the PGC (www.med.unc.edu/pgc/) in order to investigate the specificity of our findings for ASD and ADHD genetic overlap (Table S1).

Detailed sample descriptions are available in Table 1, Table S1 and the Supplementary Information.

Linkage Disequilibrium Score (LDSC) regression and correlation

Single Nucleotide Polymorphism-heritability (SNP-h^2) was estimated for EA and general intelligence on the observed scale and for psychiatric disorder samples on the liability scale using Linkage Disequilibrium Score (LDSC) regression[41] (Supplementary Information, Table S2). In addition, unconstrained LDSC correlation[23] analyses were conducted to estimate genetic correlations (r_g) among psychiatric disorders, EA and general intelligence (Supplementary Information, Table S3-S5).

Multivariable regression analysis

We applied multivariable regression (MVR) analyses (i.e. weighted multiple linear regression) to study ASD-specific and ADHD-specific associations with EA conditional on ADHD/ASD cross-disorder effects, using summary statistics. This methodology allows us to disentangle genetic relationships between traits and controls for collider bias[42] (Supplementary Information).

Genetic variant selection: Multiple ASD-related and ADHD-related variant sets were selected from ASD(iPSYCH,woADHD) and ADHD(iPSYCH) GWAS summary statistics respectively, using a range of *P*-value thresholds (Supplementary Information, $5x10^{-8} \leq P_{thr} \leq 0.5$). Throughout, MVR results are presented for two selection thresholds only: (i) $P_{thr}<0.0015$, consistent with guidelines for validating genetic instrument strength (F-statistic<10)[43] and conservative selection thresholds suggested for polygenic scoring approaches[44]. (ii) $P_{thr}<0.05$, to increase statistical power and precision of the estimates. All genetic variant sets were restricted to common (MAF>0.01), independent (PLINK[45] clumping: LD-r^2<0.25, ±500 kb) and well imputed (INFO[46]>0.7) SNPs.

Estimation of ASD-specific, ADHD-specific and cross-disorder genetic effects on EA: For each set of genetic variants, SNP estimates were extracted from ASD(iPSYCH,woADHD), ADHD and EA GWAS summary statistics (Table 1). SNP estimates for ASD variant sets were studied in ASD-MVR models (Figure 1a), SNP estimates for ADHD variant sets in ADHD-MVR models (Figure 1b). MVRs were fitted for each variant set as follows: SNP estimates for EA were simultaneously regressed on SNP estimates for ADHD and ASD and weighted by the inverse variance of SNP estimates for EA, using ordinary least square regression[47](R:stats library, Rv3.2.0). Extracted SNP estimates for ASD and ADHD were included as lnOR. Thus, ASD-specific associations with EA (ASD-MVR β_{ASD}) were estimated using ASD SNP estimates for ASD-related risk alleles; ADHD-specific associations with EA (ADHD-MVR β_{ADHD}) with ADHD SNP estimates for ADHD-related risk alleles. ADHD cross-disorder associations with EA were estimated using ADHD SNP estimates for ASD-

related alleles (ASD-MVR $\beta_{\otimes ADHD}$), ASD cross-disorder associations with EA with ASD SNP estimates for ADHD-related alleles (ADHD-MVR $\beta_{\otimes ASD}$). These MVR cross-disorder effects may capture genetic confounding, mediating effects and/or biological pleiotropy (Supplementary Information). Model fit of each MVR was compared to a single linear weighted regression model using likelihood-ratio tests (Supplementary Information).

To allow for association between the selected genetic variants and EA other than through ASD or ADHD liability, all MVR models were fitted unconstrained. As MVR estimates are thus sensitive to allelic alignment, models were fitted twice: (i) with SNP effects aligned to the risk-increasing allele for the disorder used for variant selection (i.e. ASD risk in ASD-MVRs and ADHD risk in ADHD-MVRs); (ii) a subset of variants from (i) that have the same risk allele for both ASD and ADHD (concordant variant sets, Supplementary Information).

ASD-MVRs and ADHD-MVRs were repeated using ASD(PGC) SNP estimates instead of ASD(iPSYCH,woADHD) SNP estimates.

Follow-up analyses on general intelligence: ASD-MVR and ADHD-MVR models were re-analysed using summary statistics for general intelligence, instead of EA, as described above.

Follow-up analyses of cross-disorder effects: To study cross-disorder effects on EA in detail, we restricted ASD and ADHD variant sets to SNPs that were tagged by both disorders and repeated MVR modelling (Supplementary Information). Additionally, the specificity of cross-disorder effects on EA, as detected by full ASD and ADHD variant sets, was investigated with respect to MDD, SCZ or BD SNP estimates using GWAS summary statistics (Table S1).

We applied a Bonferroni-corrected multiple testing threshold of 0.005 accounting for ten independent MVR models.

Structural equation modelling

Hypothetical structural equation models of an underlying multi-factorial model describing the interrelationship between EA, ASD and ADHD were outlined according to

theory, following a Cholesky decomposition of variance[48] (Supplementary Information). In short, a Cholesky decomposition describes a saturated model that involves the decomposition of genetic and residual variances into as many latent factors as there are observed variables[48]. Hypothetical factor loadings were derived (but not fitted) with structural equations and LDSC SNP-h^2 and r_g estimates, using EA, ASD(iPSYCH), ASD(PGC) and ADHD(iPSYCH) GWAS summary statistics (Supplementary Information, Table S3-S4). Simulations were conducted to confirm the plausibility of such a model (Supplementary Information).

Results

Multivariable regression model fitting

Using ASD-MVR, we modelled ASD-specific links with EA conditional on associations shared with ADHD, based on ASD risk alleles (Figure 1a). Applying an analogous approach, we interrogated ADHD-specific influences on EA conditional on ASD-related cross-disorder associations, as captured by ADHD risk alleles in ADHD-MVR models (Figure 1b). Compared to single linear weighted regressions modelling disorder-specific effects only, MVRs revealed a better model fit and explained ≤3% more variation in genetically predictable EA without evidence for multi-collinearity (Table S6,S7,S10,S11,S17). However, neither model allowed for causal inferences as across most MVRs regression intercepts differed from zero (Tables S6-S14,S17).

Multivariable regression analyses of EA on ASD and ADHD

Studying ASD-related variants (P_{thr}<0.0015, N_{SNPs}=1973) as part of ASD-MVR models (Figure 1a), we observed an 0.009 increase in years of schooling per log-odds in ASD-liability (ASD-MVR β_{ASD}=0.009(SE=0.003), *P*=0.002), conditional on genetic effects shared with ADHD liability (ASD-MVR $\beta_{\otimes ADHD}$=-0.029(SE=0.004), $P<1x10^{-10}$)(Figure 1c, Table S6). These ADHD cross-disorder influences on EA were larger in absolute size and had, even

though modeled with the same ASD-related alleles, an opposite direction of effect (Figure 1c,1d).

Applying an analogous approach with ADHD variant sets (ADHD-MVR, Figure 1b) revealed a complementary association profile. There was an inverse ADHD-specific association between polygenic ADHD risk and EA, while conditionally shared genetic variation with ASD showed an opposite, here positive, effect on EA (Figure 1c,1e). Using ADHD-related variants (P_{thr}<0.0015, N_{SNPs}=2717), this corresponds to a 0.012 decrease in years of education per log odds in ADHD liability (ADHD-MVR β_{ADHD}=-0.012(SE=0.003), P=4x10^{-5}), conditional on genetic effects shared with ASD (ADHD-MVR $\beta_{\otimes ASD}$=0.022(SE=0.003), P<1x10^{-10})(Figure 1c, Table S6). Increasing the number of variants in ASD-MVRs and ADHD-MVRs (P_{thr}<0.05) boosted the statistical power (Figure 1c, Table S6).

Sensitivity analyses using concordant variants, with risk-increasing effects on both ASD and ADHD risk, comprised ~80% SNPs of the initial sets and confirmed these findings. It demonstrated that observed MVR effects are (i) independent of allelic alignment (Table S7) and (ii) cannot be solely attributed to disorder-specific effects. The modelled bivariate relationships between ASD, ADHD and EA, are plotted in Figure S1 (P_{thr}<0.05).

Evidence for ASD-specific, ADHD-specific and cross-disorder associations with EA was observed for variant sets selected at different thresholds (5x10^{-8}<P_{thr}<0.5), using both ASD-MVR and ADHD-MVR, and already detectable at P_{thr}<0.0005 (Table S8,S9). Importantly, the profile of opposite cross-disorder associations with EA was replicated with MVR using SNP estimates from ASD(PGC), instead of ASD(iPSYCH,woADHD)(Table S10,S11) at P_{thr}<0.05. Thus, despite known zero genetic correlations between ADHD(iPSYCH) and ASD(PGC) (Table S3), we observed strong evidence for genetic associations (P<1x10^{-10}) between each disorder and EA using the same set of SNPs (Table S10-S11), with opposite direction of effect. MVR models including general intelligence as outcome, instead of EA, confirmed association patterns throughout (Table S12), consistent with known genetic correlations (Table S5).

Identification of cross-disorder loci

Next, we aimed to identify the largest SNP effects driving cross-disorder associations. For this, we assessed the overlap between ASD and ADHD variant sets at P_{thr}<0.0015 (N_{SNPs}≤2717) in iPSYCH samples. At this threshold, ≤4.2% of variants were cross-tagged reciprocally (LD-r^2=0.6, 500 kb), and ≤46.3% of each set were tagged when cross-disorder variant selection criteria were relaxed (P_{thr}<0.5)(Figure 2a). Conducting MVRs using tagged SNPs only (Table S13), we observed an increase in MVR effects, compared to the full set, especially when cross-tagging was performed at stringent thresholds (Figure 2b, Table S14,S15). Both ASD-MVR and ADHD-MVR models conducted with variants cross-tagged at P_{thr}<0.0015 (Table S13) revealed larger and stronger effects than MVR models involving variants that were cross-tagged at P_{thr}<0.05 (Figure 2b, Table S14,S15). The corresponding set of 83 loci (99% with concordant effects) was identified using both ASD-MVR and ADHD-MVR models. It captured 4.2% and 3.1% of the full ASD- and ADHD variant set at P_{thr}<0.0015 respectively (Figure 2a) and comprised identical or tagged proxy SNPs (Table S13). There was little evidence that any other randomly selected subset of SNPs of equal size yielded MVR effects of the same strength and magnitude (Table S16), suggesting effect specificity. These SNPs mapped to 52 genes only (Figure S2) that were most strongly enriched for phosphatidylinositol-3-lipid-kinase (PI3K) pathways (Table S17) and contained multiple regulatory RNAs (Figure S2).

Specificity of ADHD/ASD cross-disorder genetic effects

To assess the specificity of opposite cross-disorder associations, as captured by ASD and ADHD variants respectively, we also studied cross-disorder effects using MDD, SCZ and BD SNP estimates. We identified similar patterns for several combinations of disorders, predominantly at P_{thr}<0.05 (Figure 3, Table S18), that were consistent with known genetic correlations (Table S3-S4). For ASD-MVR, with reference to positive ASD-specific effects on EA, opposite patterns were detected in combination with MDD (e.g. P_{thr}<0.05: ASD-MVR

$\beta_{\otimes MDD}$=-0.012, SE=0.001, P<1x10^{-10}). For ADHD-MVR, with respect to inverse ADHD-specific effects on EA, opposite MVR effects were found for BD (e.g. P_{thr}<0.05: ADHD-MVR $\beta_{\otimes BD}$=0.008, SE=0.001, P<1x10^{-10}). However, these cross-disorder associations were considerably smaller than cross-disorder effects for both ADHD and ASD within iPSYCH (Figure 1c,3).

Multi-factor model of genetic interrelations between ASD, ADHD and educational attainment

To integrate MVR findings with known genetic interrelationships between EA, ASD and ADHD we translated the latter into path coefficients for a saturated structural equation model (Figure 4). Consistent with a Cholesky decomposition (Supplementary Information, Figure S3), there are at least two sources of shared genetic variation between ASD and ADHD. The first genetic factor (A1, EA/ADHD/ASD) captures shared variation between EA, ASD and ADHD, and predicts negative genetic covariance between ASD and ADHD, consistent with MVR findings in this study. The second genetic factor (A2, ADHD/ASD) acts independently of A1 and explains positive genetic covariance between ASD and ADHD, reflecting known positive or null genetic correlations between disorders [21,22]. The third genetic factor (A3) allows for ASD-specific (Figure 4) or ADHD-specific (Figure S4) variation, dependent on the definition of A2 (Supplementary Information). The observed net covariance between ASD and ADHD reflects thus the sum of negative and positive covariance contributions. Consequently, ASD/ADHD genetic overlap might be reduced, as hypothesised for ASD(iPSYCH)/ADHD(iPSYCH) (Figure 4a,S4a). It might also be completely abolished, as hypothesised for ASD(PGC)/ADHD(iPSYCH) (Figure 4b,S4b), supported by simulations (Supplementary Information, Table S19).

The assumption of EA-related negative genetic covariance between ASD and ADHD finds also support through LDSC correlation analyses. Genetic correlations with EA for ASD(iPSYCH,woADHD) (r_g=0.23(SE=0.03), P=<1x10^{-10}), excluding comorbid ADHD patients, exceeded those for the full ASD(iPSYCH) sample (r_g=0.16(SE=0.03), P=4x10^{-7}), although 95%-confidence intervals overlap. In contrast, inverse genetic correlations between

ADHD and EA (r_g=-0.49(SE=0.03), *P*=<1x10^{-10}) were dampened, when ADHD(iPSYCH) and ASD(PGC) summary statistics were combined (r_g=-0.33(SE=0.03), *P*=<1x10^{-10})(Figure S5).

Discussion

This study provides strong and consistent evidence that EA-related polygenic variation is shared across ASD and ADHD. Here, we show that different combinations of the same risk-increasing alleles can result in ASD-related positive and ADHD-related negative association profiles with genetically predictable EA. This suggests the presence of pleiotropic mechanisms, where opposite association profiles with EA can be encoded across the same polygenic sites, without involving distinct sets of SNPs or genes, potentially leading to negative correlations.

The pattern of ASD- and ADHD-specific effects on EA, in combination with opposite cross-disorder associations, was (i) reciprocally detectable using both ASD and ADHD-related variant sets, (ii) replicated using ASD(PGC) summary statistics, (iii) independent of risk allele alignment and (iv) consistent with the previously reported genetic overlap between EA, ASD and ADHD[21,29]. The detection of opposite association profiles that strengthened with increasing numbers of SNPs confirmed that pleiotropic effects are detectable for the majority of trait-associated variants in the genome[49]. Importantly, the strength and absolute size of cross-disorder effects was comparable to disorder-specific influences on EA, i.e. ASD-specific and ADHD-specific effects based on same variant set, or larger. Similar patterns were also present for the same disorder across different MVRs. This underlines that cross-disorder associations are driven by a substantial proportion of subthreshold variants that are associated with both ASD and ADHD. However, against this shared polygenic background involving several thousands of variants, we also identified ~80 loci (~50 genes) that exerted discernably larger MVR effects when signals were followed-up in the powerful iPSYCH samples.

The identification of opposite EA-related polygenic cross-disorder associations for ADHD and ASD that involve different combinations of the same risk alleles may relate to different mechanisms. First, there is mounting evidence that ASD and ADHD share some underlying aetiological mechanisms. For example, CNVs in ASD and ADHD indicate similar biological pathways[19] and both disorders carry a similar burden of rare protein-truncating variants, implicating many shared genes[50]. Additionally, mice carrying homozygous mutations in Shank2, an ASD high-risk locus encoding a synaptic protein, display both extreme hyperactive and autistic-like behaviour[51]. Despite these biological commonalties, the assignment of clinical diagnoses to patients comorbid for ASD and ADHD symptomatology has been, until the introduction of the Diagnostic Statistical Manual of Mental Disorders 5th edition (DSM-5)[52], less formalised. Participants in large GWASs have been predominantly diagnosed with previous classification systems[53,54], where hierarchical diagnostic criteria did not allow for a diagnosis of ADHD when symptoms occurred during the course of a pervasive developmental disorder. Furthermore, comorbid symptoms within clinical ASD and ADHD often occur at the subthreshold level[3]. This suggests that patients with comorbid ASD and ADHD symptoms might have been assigned to either diagnostic entity, depending on the symptoms that presented first. Thus, genetic factors underlying this shared ASD and ADHD symptomatology might be either analysed within the context of an ASD or an ADHD genetic architecture, potentially exacerbating genetic similarities.

Second, shared ADHD/ASD cross-disorder variation with opposite direction of effects may involve epistasis[55], such that ASD-specific and ADHD-specific genetic factors may shape the direction and effect of ADHD/ASD cross-disorder effects on EA. For example, following an omnigenic model[56,57], disorder-specific 'peripheral' genetic influences would control shared ADHD/ASD cross-disorder 'core' variation.

Third, ADHD/ASD cross-disorder variation may involve high plasticity genes, exerting different effects within differing environments. The strongest signals driving the observed opposite cross-disorder effects in iPSYCH samples were found near several miRNA and

lncRNA loci that can be influenced by environmental signals[58]. They were also enriched for PI3K and linked genes with catalytic and regulatory roles in normal cell function and cancer[59]. Thus, the symptoms and behavioural spectrum of an individual at high genetic risk for both ASD and ADHD may differ based on the exposure to different home environment (e.g. household income, neighbourhood SES and child age when mother returned to work). This is consistent with recent findings for depression reporting a modulatory effect of stressors on genes manifesting as an environment-induced development of depression[60].

Opposite cross-disorder associations implicating the same alleles, although in different combinations, may potentially lead to inverse EA-related genetic covariance between ASD and ADHD that reduces net genetic overlap. This notion is supported by simulations and LDSC-based genetic correlation analyses. The discovered inverse cross-disorder effects on EA were stronger and larger for ADHD and ASD compared to cross-disorder effects involving other psychiatric disorders. However, they are unlikely to be limited to polygenic ASD and ADHD risk and inverse genetic overlap may affect also other cross-disorder investigations.

Our results are strengthened, as we replicate evidence for ADHD/ASD cross-disorder associations using two independent ASD collections, iPSYCH and PGC. Therefore, our findings are robust across diagnostic classification systems for clinical ASD, routes of patient ascertainment, and association analysis designs[36,37]. A weakness is that our results could be affected by presentation bias, such that children with ASD might be more often labelled with ADHD, due to a higher proportion of ADHD symptoms in ASD[61]. Furthermore, controls are shared across iPSYCH GWAS summary statistics, potentially leading to inflated type I error[62,63]. This is, however, unlikely, given the opposite direction of observed effects and replication within the independent ASD(PGC) sample. As symptom heterogeneity may affect genetic overlap between neurodevelopmental disorders, EA and cognition-related traits[21,64], future studies with access to this information will be required to fully understand the underlying complex multivariate interrelations.

Conclusion

We show that EA-related polygenic variation is shared across ASD and ADHD genetic architectures, and that different combinations of the same risk alleles can encode opposite disorder-specific association profiles with EA. Our findings are consistent with inverse EA-related genetic relationships between ASD and ADHD that may contribute to the net genetic overlap between disorders.

Acknowledgements

This research was facilitated by the Social Science Genetic Association Consortium (SSGAC), Complex Trait Genetics (CTG) lab, Psychiatric Genomics Consortium (PGC) and iPSYCH-Broad-PGC ASD Consortium, by providing access to genome-wide summary statistics.

This publication is the work of the authors and EV and BSTP will serve as guarantors for the contents of this paper. EV and BSTP are supported by the Max Planck Society. BSTP is supported by the Simons Foundation (Award ID: 514787). The iPSYCH project is funded by the Lundbeck Foundation (grant no R102-A9118 and R155-2014-1724) and the universities and university hospitals of Aarhus and Copenhagen. ADB is also supported by the EU's Horizon 2020 programme (grant no 667302, CoCA). Data handling and analysis was supported by NIMH (1U01MH109514-01 to Michael O'Donovan and Anders D Børglum). High-performance computer capacity for handling and statistical analysis of iPSYCH data on the GenomeDK HPC facility was provided by the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark (grant to Anders D Børglum) and Center for Genomics and Personalized Medicine, Aarhus, Denmark. We thank Simon E. Fisher for helpful discussions of the mansucript.

Conflict of interest

The authors declare no conflict of interest.

References

1 Thapar A, Cooper M. Attention deficit hyperactivity disorder. *The Lancet* 2016; **387**: 1240–1250.

2 Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. *The Lancet* 2018; **392**: 508–520.

3 Antshel KM, Zhang-James Y, Wagner KE, Ledesma A, Faraone SV. An update on the comorbidity of ADHD and ASD: a focus on clinical management. *Expert Rev Neurother* 2016; **16**: 279–293.

4 Stergiakouli E, Hamshere M, Holmans P, Langley K, Zaharieva I, deCODE Genetics *et al.* Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am J Psychiatry* 2012; **169**: 186–194.

5 Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB *et al.* Most genetic risk for autism resides with common variation. *Nature Genetics* 2014; **46**: 881–885.

6 Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, Witherspoon K *et al.* Excess of rare, inherited truncating mutations in autism. *Nat Genet* 2015; **47**: 582–588.

7 Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R *et al.* Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet* 2010; **376**: 1401–1408.

8 Rommelse NNJ, Franke B, Geurts HM, Hartman CA, Buitelaar JK. Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur Child Adolesc Psychiatry* 2010; **19**: 281–295.

9 Leitner Y. The co-occurrence of autism and attention deficit hyperactivity disorder in children - what do we know? *Front Hum Neurosci* 2014; **8**: 268.

10 Ronald A, Simonoff E, Kuntsi J, Asherson P, Plomin R. Evidence for overlapping genetic influences on autistic and ADHD behaviours in a community twin sample. *J Child Psychol Psychiatry* 2008; **49**: 535–542.

11 Stergiakouli E, Davey Smith G, Martin J, Skuse DH, Viechtbauer W, Ring SM *et al.* Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Molecular Autism* 2017; **8**: 18.

12 Ronald A, Larsson H, Anckarsäter H, Lichtenstein P. Symptoms of autism and ADHD: A Swedish twin study examining their overlap. *Journal of Abnormal Psychology* 2014; **123**: 440–451.

13 Polderman TJC, Hoekstra RA, Posthuma D, Larsson H. The co-occurrence of autistic and ADHD dimensions in adults: an etiological study in 17□770 twins. *Transl Psychiatry* 2014; **4**: e435.

14 Taylor MJ, Charman T, Ronald A. Where are the strongest associations between autistic traits and traits of ADHD? evidence from a community-based twin study. *Eur Child Adolesc Psychiatry* 2015; **24**: 1129–1138.

15 Taylor MJ, Charman T, Robinson EB, Plomin R, Happé F, Asherson P *et al.* Developmental associations between traits of autism spectrum disorder and attention deficit hyperactivity disorder: a genetically informative, longitudinal twin study. *Psychol Med* 2013; **43**: 1735–1746.

16 Martin J, Hamshere ML, Stergiakouli E, O'Donovan MC, Thapar A. Genetic Risk for Attention-Deficit/Hyperactivity Disorder Contributes to Neurodevelopmental Traits in the General Population. *Biol Psychiatry* 2014; **76**: 664–671.

17 Lichtenstein P, Carlström E, Råstam M, Gillberg C, Anckarsäter H. The Genetics of Autism Spectrum Disorders and Related Neuropsychiatric Disorders in Childhood. *AJP* 2010; **167**: 1357–1363.

18 Ghirardi L, Brikell I, Kuja-Halkola R, Freitag CM, Franke B, Asherson P *et al.* The familial co-aggregation of ASD and ADHD: a register-based cohort study. *Mol Psychiatry* 2018; **23**: 257–262.

19 Martin J, Cooper M, Hamshere ML, Pocklington A, Scherer SW, Kent L *et al.* Biological overlap of attention-deficit/hyperactivity disorder and autism spectrum disorder: evidence from copy number variants. *J Am Acad Child Adolesc Psychiatry* 2014; **53**: 761-770.e26.

20 Martin J, Taylor MJ, Lichtenstein P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol Med* 2018; **48**: 1759–1774.

21 Grove J, Ripke S, Als TD, Mattheisen M, Walters R, Won H *et al.* Common risk variants identified in autism spectrum disorder. *bioRxiv* 2017. doi:10.1101/224774.

22 Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.

23 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**: 1236–41.

24 Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016; **533**: 539–42.

25 Martin J, Hamshere ML, Stergiakouli E, O'Donovan MC, Thapar A. Neurocognitive abilities in the general population and composite genetic risk scores for attention-deficit hyperactivity disorder. *J Child Psychol Psychiatry* 2015; **56**: 648–656.

26 Consortium TB, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J *et al.* Analysis of shared heritability in common disorders of the brain. *Science* 2018; **360**: eaap8757.

27 Stergiakouli E, Martin J, Hamshere ML, Heron J, St Pourcain B, Timpson NJ *et al.* Association between polygenic risk scores for attention-deficit hyperactivity disorder and educational and cognitive outcomes in the general population. *Int J Epidemiol* 2017; **46**: 421–428.

28 Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, Leeuw CA de *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics* 2018; **50**: 912–919.

29 Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 2018; : 1.

30 Clarke T-K, Lupton MK, Fernandez-Pujals AM, Starr J, Davies G, Cox S *et al.* Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. *Mol Psychiatry* 2016; **21**: 419–425.

31 Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J *et al.* Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat Genet* 2017; **49**: 978–985.

32 Russell G, Ford T, Rosenberg R, Kelly S. The association of attention deficit hyperactivity disorder with socioeconomic disadvantage: alternative explanations and evidence. *J Child Psychol Psychiatr* 2014; **55**: 436–445.

33 Durkin MS, Maenner MJ, Meaney FJ, Levy SE, DiGuiseppi C, Nicholas JS *et al.* Socioeconomic Inequality in the Prevalence of Autism Spectrum Disorder: Evidence from a U.S. Cross-Sectional Study. *PLOS ONE* 2010; **5**: e11551.

34 Rommelse NNJ, Geurts HM, Franke B, Buitelaar JK, Hartman CA. A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience & Biobehavioral Reviews* 2011; **35**: 1363–1396.

35 Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* 2018; **50**: 1112–1121.

36 Pedersen CB, Bybjerg-Grauholm J, Pedersen MG, Grove J, Agerbo E, Bækvad-Hansen M *et al.* The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol Psychiatry* 2017. doi:10.1038/mp.2017.196.

37 Autism Spectrum Disorder Working Group of the Psychiatry Genomics Consortium. Dataset: PGC-ASD summary statistics from a meta-analysis of 5,305 ASD-diagnosed cases and 5,305 pseudocontrols of European descent (based on similarity to CEPH reference genotypes). 2015.http://www.med.unc.edu/pgc/results-and-downloads.

38 Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010; **11**: 288.

39 Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* 2018; **50**: 668.

40 Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW *et al.* Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* 2018; **173**: 1705-1715.e16.

41 Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* 2015; **47**: 291–5.

42 Aschard H, Vilhjálmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet* 2015; **96**: 329–339.

43 Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol* 2011; **40**: 740–752.

44 Wray NR, Lee SH, Mehta D, Vinkhuyzen AAE, Dudbridge F, Middeldorp CM. Research Review: Polygenic methods and their application to psychiatric traits. *J Child Psychol Psychiatr* 2014; **55**: 1068–1087.

45 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007; **81**: 559–575.

46 Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLOS Genetics* 2009; **5**: e1000529.

47 Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting Causal Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age at Menarche and Risk of Breast Cancer. *Genetics* 2017; : genetics.300191.2017.

48 Neale M, Boker S, Xie G, Maes HHM. *Mx: Statistical modeling*. 7th ed. Department of Psychiatry: Richmond, 2006.

49 Watanabe K, Stringer S, Frei O, Mirkov MU, Polderman TJC, Sluis S van der *et al.* A global view of pleiotropy and genetic architecture in complex traits. *bioRxiv* 2018; : 500090.

50 Satterstrom FK, Walters RK, Singh T, Wigdor EM, Lescai F, Demontis D *et al.* ASD and ADHD have a similar burden of rare protein-truncating variants. *bioRxiv* 2018; : 277707.

51 Schmeisser MJ, Ey E, Wegener S, Bockmann J, Stempel AV, Kuebler A *et al.* Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* 2012; **486**: 256–260.

52 Association AP. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5*. 5 edition. American Psychiatric Publishing: Washington, D.C, 2013.

53 *International Statistical Classification of Diseases and Related Health Problems*. World Health Organization: Malta, 2010.

54 Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 1994; **24**: 659–685.

55 Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 2002; **11**: 2463–2468.

56 Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 2017; **169**: 1177–1186.

57 Liu X, Li YI, Pritchard JK. Trans effects on gene expression can drive omnigenic inheritance. *bioRxiv* 2018; : 425108.

58 Morris KV, Mattick JS. The rise of regulatory RNA. *Nature Reviews Genetics* 2014; **15**: 423–437.

59 Fruman DA, Chiu H, Hopkins BD, Bagrodia S, Cantley LC, Abraham RT. The PI3K Pathway in Human Disease. *Cell* 2017; **170**: 605–635.

60 Gonda X, Hullam G, Antal P, Eszlari N, Petschner P, Hökfelt TG *et al.* Significance of risk polymorphisms for depression depends on stress exposure. *Scientific Reports* 2018; **8**: 3946.

61 Mayes SD, Calhoun SL, Mayes RD, Molitoris S. Autism and ADHD: Overlapping and discriminating symptoms. *Research in Autism Spectrum Disorders* 2012; **6**: 277–285.

62 Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two□sample Mendelian randomization. *Genet Epidemiol* 2016; **40**: 597–608.

63 Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet* 2013; **14**: 507–515.

64 Greven CU, Harlaar N, Dale PS, Plomin R. Genetic Overlap between ADHD Symptoms and Reading is largely Driven by Inattentiveness rather than Hyperactivity-Impulsivity. *J Can Acad Child Adolesc Psychiatry* 2011; **20**: 6–14.

Tables

Table 1: Sample description

Source	Phenotype	Consortium	GWAS	Imputation reference panel	N	Analyses
Clinical sample	ASD	iPSYCH	ASD(iPSYCH)	1000 Genomes phase 3	35,740 (13,076 cases)	LDSC
			ASD(iPSYCH,woADHD)	1000 Genomes phase 3	32,985 (10,321 cases)	LDSC, MVR
		PGC	ASD(PGC)	1000 Genomes phase 1 (v3)	10,610 (5,305 cases)	LDSC, MVR
	ADHD	iPSYCH	ADHD	1000 Genomes phase 3	37,076 (14,584 cases)	LDSC, MVR
Population sample	Years of schooling	SSGAC	EA	1000 Genomes phase 3[1]	766,345	LDSC, MVR

1. Predominantly 1000 genomes phase 3, see Lee et al.[35]

Abbreviations: ASD, Autism Spectrum Disorder; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; SSGAC, Social Science Genetic Consortium; LDSC, Linkage Disequilibrium Score; MVR, multivariable regression; woADHD; without ADHD.

All individuals were of European descent.

Figure legends

Figure 1: ASD-specific, ADHD-specific and cross-disorder effects on EA

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; MVR, multivariable regression; P_{thr}, P-value threshold.

Sets of independent genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS summary statistics at different *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). Corresponding SNP estimates for ASD, ADHD and EA were subsequently extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS summary statistics respectively. **(a)** Schematic ASD-MVR model estimating ASD-specific (ASD-MVR β_{ASD}) and ADHD cross-disorder (ASD-MVR $\beta_{\otimes ADHD}$) effects on EA with ASD variant sets. ASD-MVR β_{ASD} were estimated with ASD variant sets and corresponding ASD SNP estimates and quantify here the change in years of schooling per log-odds ASD liability. ADHD cross-disorder effects (ASD-MVR $\beta_{\otimes ADHD}$) were assessed with ADHD SNP estimates for ASD variant sets and quantify the change in years of schooling per log-odds ADHD liability. **(b)** Schematic ADHD-MVR model estimating ADHD-specific (ADHD-MVR β_{ADHD}) and ASD cross-disorder effects (ADHD-MVR $\beta_{\otimes ASD}$) on EA using ADHD variant sets. ADHD-MVR β_{ADHD} were estimated with ADHD variant sets and corresponding ADHD SNP estimates and estimate the change in years of schooling per log-odds ADHD liability. ASD cross-disorder effects (ADHD-MVR $\beta_{\otimes ASD}$) were assessed with ASD SNP estimates for ADHD variant sets and estimates quantify the change in years of schooling per log-odds ADHD liability. **(c)** ASD-specific (ASD-MVR β_{ASD}), ADHD-specific (ADHD-MVR β_{ADHD}) and cross-disorder effects (ASD-MVR $\beta_{\otimes ADHD}$, ADHD-MVR $\beta_{\otimes ASD}$) were estimated aligning SNP effects according to ASD risk (ASD variant set) and ADHD risk (ADHD variant set) respectively. All MVR effects are presented with respect to years of schooling, i.e. per increase in log odds of ASD or ADHD liability respectively. Bars represent 95% confidence intervals. **(d)** 3D scatter plot of ASD SNP estimates (x-axis), ADHD SNP estimates (y-axis) and EA SNP estimates (z-axis) for ASD-related variants (P_{thr}<0.0015). The multivariable regression plane reflects ASD-specific and ADHD cross-disorder associations, as shown in Figure 1c. **(e)** 3D scatter plot of ASD SNP estimates (x-axis), ADHD SNP estimates (y-axis) and EA SNP estimates (z-axis) for ADHD-related variants (P_{thr}<0.0015). The multivariable regression plane reflects ADHD-specific and ASD cross-disorder associations, as shown in Figure 1c.

Figure 2: ASD-specific, ADHD-specific and cross-disorder effects on educational attainment using ASD and ADHD tagSNPs (P_{thr}<0.0015)

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; MVR, multivariable regression; P_{thr}, P-value threshold.

(a) Percentage of ASD variants (P_{thr}<0.0015) cross-tagged by ADHD variants at various *P*-value selection thresholds (within 500kb and LD-r^2≥0.6) **(b)** Percentage of ADHD variants (P_{thr}<0.0015) cross-tagged by ASD variants at various *P*-value selection thresholds (within 500kb and LD-r^2≥0.6) **(c)** ASD-MVR and ADHD-MVR analyses based on cross-tagged SNPs only. SNP estimates were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS summary statistics. ASD-MVRs were used to estimate ASD-specific (ASD-MVR β_{ASD}) and ADHD cross-disorder (ASD-MVR $\beta_{\otimes ADHD}$) effects on EA using cross-tagged subsets of ASD variants (2a). ADHD-MVRs were used to estimate ADHD-specific (ADHD-MVR β_{ADHD}) and ASD cross-disorder (ADHD-MVR $\beta_{\otimes ASD}$) effects on EA using cross-tagged subsets of ADHD variants (2a). All MVR effects, except ADHD-MVR β_{ADHD} cross-tagged at P_{thr}<0.05, passed the experiment-wide significance threshold of 0.005.

Figure 3: ASD-specific, ADHD-specific and cross-disorder effects (MDD, SCZ and BD) on educational attainment

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; MDD; Major Depressive Disorder; MVR, multivariable regression; SCZ, Schizophrenia; P_{thr}, P-value threshold.

*MVR effects passing a Bonferroni corrected *P*-value threshold of 0.005.

Sets of independent genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) summary statistics at different *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). Per variant set three MVR models were run, in addition to ASD-MVR and ADHD-MVR models (Figure 1). In ASD-MVR models (Figure 1a), ADHD SNP estimates were once replaced with MDD SNP estimates, once with SCZ SNP estimates and once with BD SNP estimates. Likewise, in ADHD-MVR models (Figure 1b) ASD SNP estimates were replaced with SNP estimates for MDD, SCZ and BD. SNP estimates were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS summary statistics. For ASD-MVR models (using ASD variant sets), ASD-specific effects (ASD-MVR β_{ASD}) are shown in addition to cross-disorder associations with EA for MDD (ASD-MVR $\beta_{\otimes MDD}$), SCZ (ASD-MVR $\beta_{\otimes SCZ}$) and BD (ASD-MVR $\beta_{\otimes MDD}$). For ADHD-MVR models (using ADHD variant sets), ADHD-specific effects (ADHD-MVR β_{ADHD}) and cross-disorder associations with EA for MDD (ADHD-MVR $\beta_{\otimes MDD}$), SCZ (ADHD-MVR $\beta_{\otimes SCZ}$) and BD (ADHD-MVR $\beta_{\otimes MDD}$) are shown. ASD-specific (ASD-MVR β_{ASD}), ADHD-specific (ADHD-MVR β_{ADHD}) and cross-disorder effects were estimated aligning SNP effects according to ASD risk (ASD variant sets) and ADHD risk (ADHD variant sets) respectively. All MVR effects are presented with respect to years of schooling, i.e. per increase in log odds liability in psychiatric disorder of interest. Bars represent 95% confidence intervals.

Figure 4: Hypothetical multi-factor model allowing for ASD-specific genetic influences

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; SNP h^2, SNP heritability; SNP r_g, SNP genetic correlation, cov_g, genetic covariance; woADHD, without ADHD

Proposed multi-factor model for EA, ADHD and ASD. The model consists of two shared genetic influences, as captured by common variants within an infinitely large population. The first genetic factor (A1, shared EA/ADHD/ASD), explains variation in EA, ADHD and ASD. It allows for a negative genetic covariance between ASD and ADHD, consistent with MVR findings in this study. The second genetic factor (A2, shared ADHD/ASD) acts independently of A1, explaining positive genetic covariance between ASD and ADHD. Factor loadings ("a") were derived from previously reported genetic LDSC SNP-heritability and genetic correlation estimates using GWAS summary statistics (grey boxes, Table S2-S3). Shared ADHD/ASD genetic influences (A2) were modelled through ADHD genetic influences including those that are shared with ASD (allowing for ASD-specific effects, A3). Observable phenotypic measures are represented by squares, while latent genetic factors are represented by circles. Single headed arrows denote genetic factor loadings ("a"), double-headed arrows genetic correlations ("r_g"). To improve clarity, residual influences are not shown and unit variances for latent variables were omitted. Each factor loading ("a") for the Cholesky decomposition of a trivariate trait is described in detail in the Supplement (Supplementary Information, Figure S3). **(a)** Multi-factor model consistent with ASD(iPSYCH), ADHD(iPSYCH) and EA(SSGAC) summary statistics. **(b)** Multi-factor model consistent with ASD(PGC), ADHD(iPSYCH) and EA(SSGAC) summary statistics and supported by simulations (Supplementary information, Table S19).

Figure 1

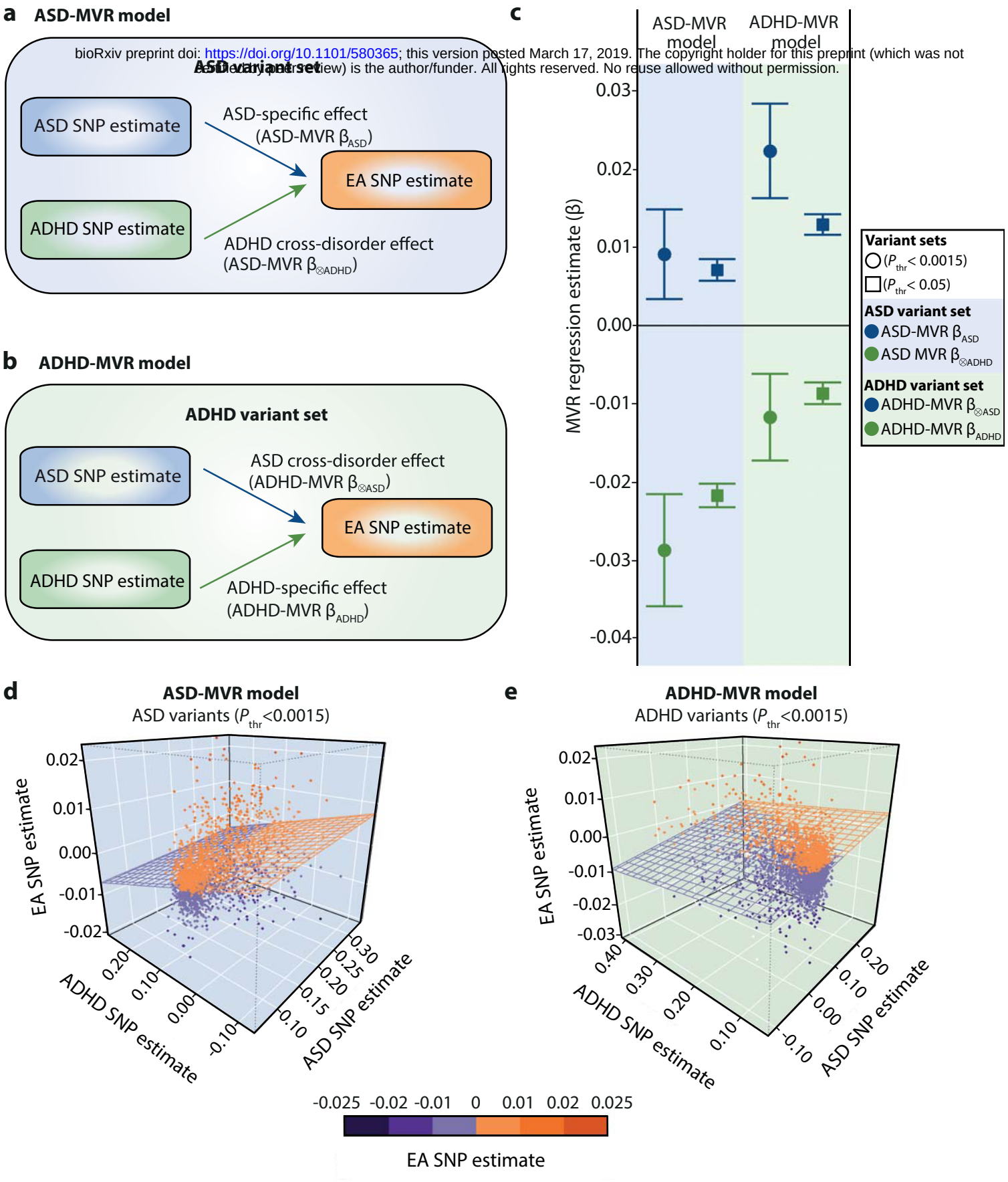
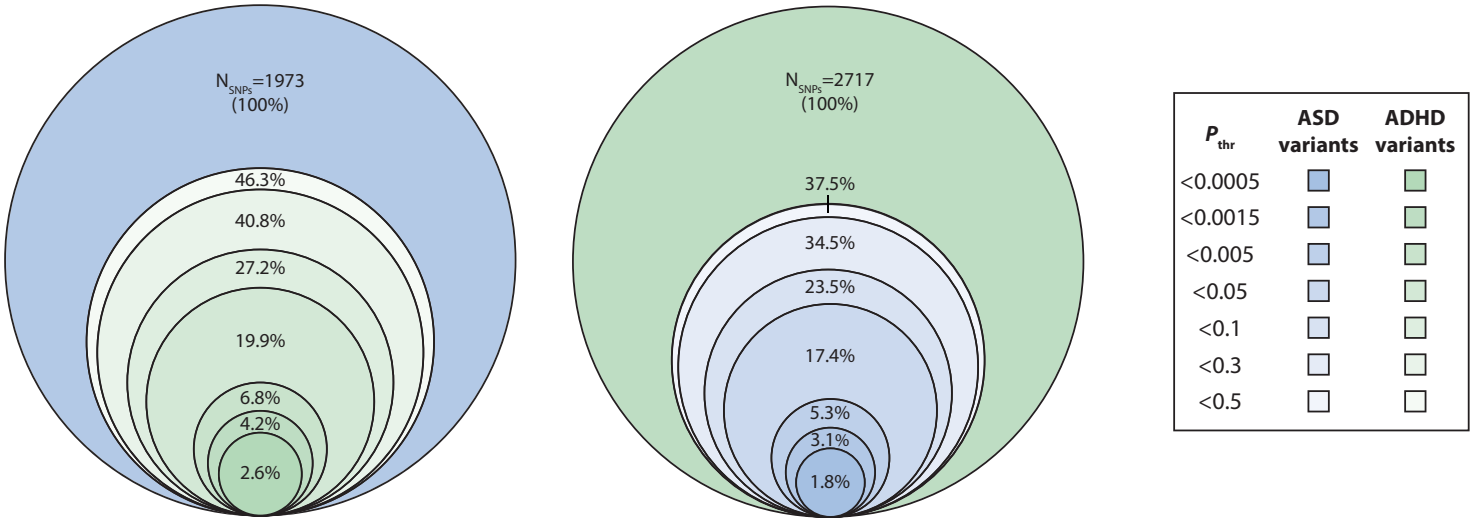


Figure 2

a



b

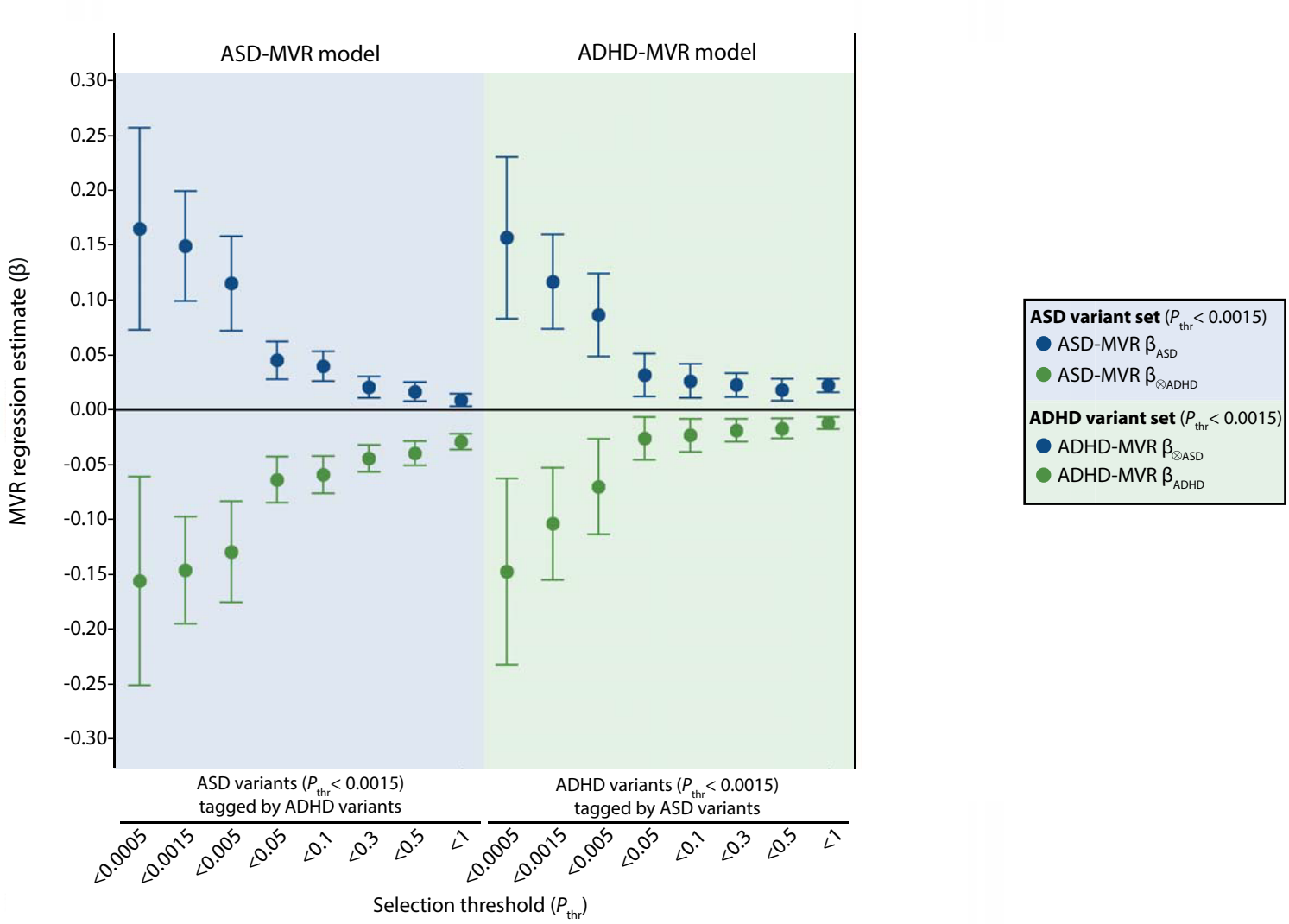


Figure 3

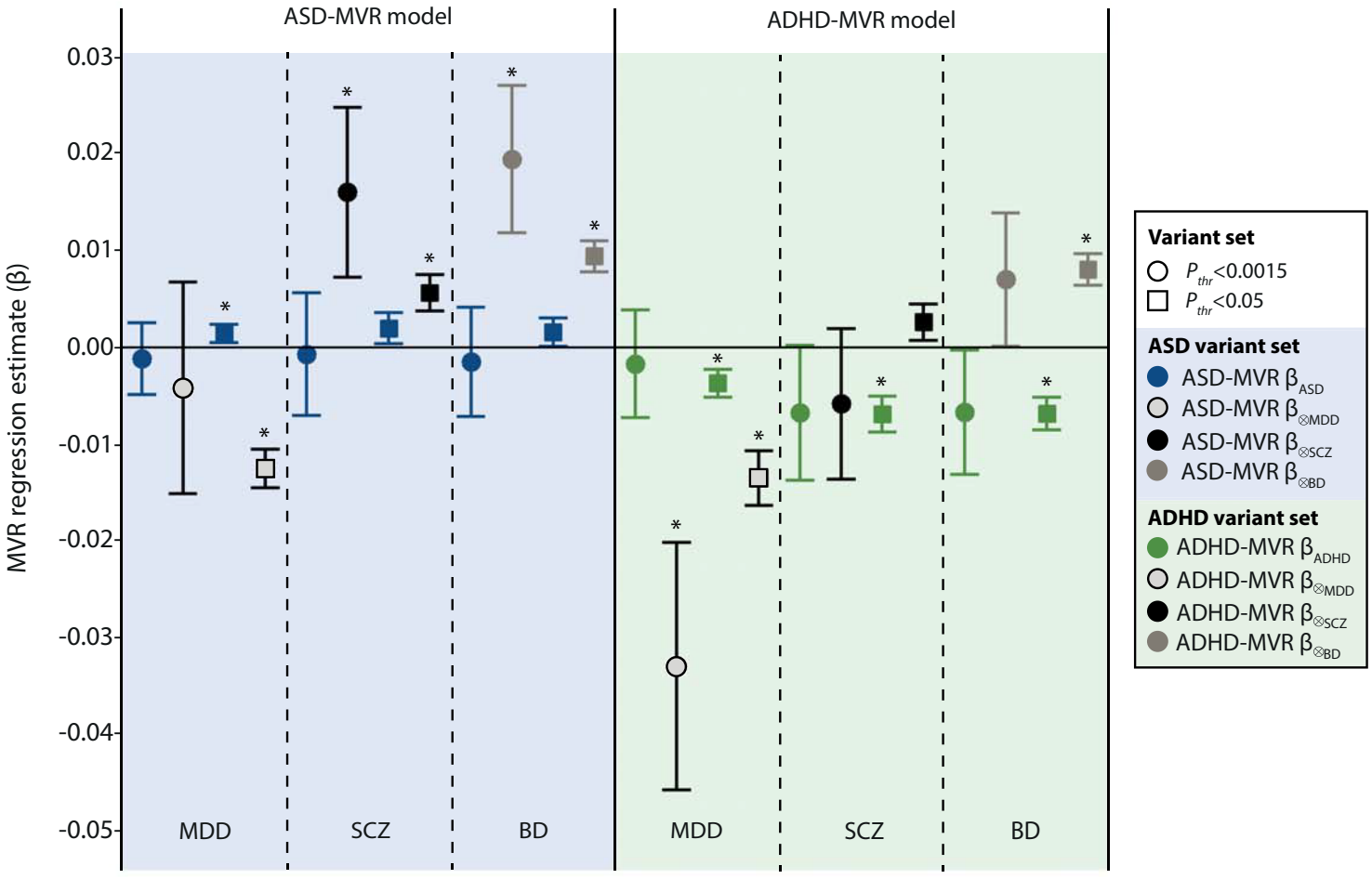


Figure 4