

Abstract word count: 194  
Main word count: 582  
Results word count: 1,441  
Discussion word count: 1,865  
Methods word count: 4,229  
# Figures: 5  
# Tables: 1  
# Extended Data Figures: 10  
# Supplemental Tables: 25  
# References: 34  
# Methods References: 30

## **Phenome-wide Investigation of Behavioral, Environmental, and Neural Associations with Cross-Disorder Genetic Liability in Youth of European Ancestry**

Sarah E. Paul, M.A.,<sup>1</sup> Sarah M.C. Colbert, B.S.,<sup>2</sup> Aaron J. Gorelik, B.S.,<sup>1</sup> Isabella S. Hansen, B.S.,<sup>1</sup> Nagella, I.,<sup>1</sup> Blaydon, L.,<sup>1</sup> Hornstein, A.,<sup>1</sup> Emma C. Johnson, Ph.D.,<sup>2</sup> Alexander S. Hatoum, Ph.D.,<sup>2</sup> David A.A. Baranger, Ph.D.,<sup>1</sup> Nourhan M. Elsayed, M.A.,<sup>1</sup> Deanna M. Barch, Ph.D.,<sup>1,2,3</sup> Ryan Bogdan, Ph.D.,<sup>1</sup> Nicole R. Karcher, Ph.D.<sup>2</sup>

<sup>1</sup> Department of Psychological and Brain Sciences, Washington University, St. Louis, MO

<sup>2</sup> Department of Psychiatry, Washington University School of Medicine, St. Louis, MO

<sup>3</sup> Department of Radiology, Washington University School of Medicine, St. Louis, MO

Author's note: Correspondence concerning this article should be addressed to Sarah E. Paul (email: [spaul24@wustl.edu](mailto:spaul24@wustl.edu), address: CB 1125 Psychological and Brain Sciences Bldg Room 451, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA

## **ABSTRACT**

Etiologic insights into psychopathology may be gained by using hypothesis-free methods to identify associations between genetic risk for broad psychopathology and phenotypes measured during adolescence, including both markers of child psychopathology and intermediate phenotypes such as neural structure that may link genetic risk with outcomes. We conducted a phenome-wide association study (phenotype n=1,269-1,694) of polygenic risk scores (PRS) for broad spectrum psychopathology (i.e., Compulsive, Psychotic, Neurodevelopmental, and Internalizing) in youth of PCA-selected European ancestry (n=5,556; ages 9-13) who completed the baseline and/or two-year follow-up of the ongoing Adolescent Brain Cognitive Development<sup>SM</sup> (ABCD) Study. We found that Neurodevelopmental and Internalizing PRS were significantly associated with a host of proximal as well as distal phenotypes (Neurodevelopmental: 187 and 211; Internalizing: 122 and 173 phenotypes at baseline and two-year follow-up, respectively), whereas Compulsive and Psychotic PRS showed zero and one significant associations, respectively, after Bonferroni correction. Neurodevelopmental PRS were further associated with brain structure metrics (e.g., total volume, mean right hemisphere cortical thickness), with only cortical volume indirectly linking Neurodevelopmental PRS to grades in school. Genetic variation influencing risk to psychopathology manifests broadly as behaviors, psychopathology symptoms, and related risk factors in middle childhood and early adolescence.

*Keywords:* Phenome-wide Association Study (PheWAS), polygenic risk scores, Neurodevelopmental, Internalizing

## MAIN

Psychopathology is common, highly comorbid, heritable, and associated with a host of negative outcomes (e.g., increased morbidity and mortality)<sup>1-3</sup>. The high polygenicity of and genetic covariance among psychiatric disorders parallel the large degree of symptom and diagnostic overlap at the phenotypic level<sup>4</sup>. Although progress has been made in characterizing the patterns of comorbidity and consolidating new and existing findings into shared frameworks of overarching functional domains (e.g., HiTOP,<sup>5</sup> RDoC<sup>6</sup>), etiological and mechanistic insight needed for prevention and mitigation efforts is still limited.

Recent advances in genomic methodologies follow decades of clinical phenotype refinement, shifting the focus from discrete diagnostic categories to higher order symptom clusters. With recently developed techniques, such as Genomic Structural Equation Modeling (Genomic SEM)<sup>7</sup>, summary statistics from Genome-wide Association Studies (GWAS) of psychiatric phenotypes can be leveraged to identify genomic variants associated with higher order dimensions of psychopathology. Both phenotypically and genetically, transdiagnostic factors offer improved validity and power to identify genomic risk loci and associations with other traits of interest, particularly as GWAS sample sizes increase and further advance the utility of Genomic SEM<sup>7,8</sup>. Polygenic scores (PRS) generated from multivariate GWAS tend to explain more variance than those from univariate GWAS, enhancing their value as proxies for overall genetic liability<sup>7</sup>. Better power and predictive utility of PRS enhances the likelihood of identifying potential intermediate phenotypes observed in childhood and adolescence that may link genetic susceptibility to psychopathology outcomes, thus improving mechanistic insight. For instance, studies have begun to investigate neural metrics as possible mechanisms driving associations between genetic liability and phenotypic expression<sup>e.g., 9,10</sup>.

Despite novel insights gained by hypothesis-driven psychopathology research with PRS (e.g., that ADHD PRS is associated with increased screen time in youth<sup>11</sup>), such work is constrained in its potential scope. Further etiological insights can be gained using hypothesis-free methods, such as phenome-wide association studies (PheWAS), which explore associations between genetic risk and hundreds or thousands of phenotypes spanning multiple domains<sup>12</sup>. Effects that survive correction for multiple testing may replicate previously discovered genotype-phenotype associations or identify novel associations that may further drive the development of new hypotheses and theory.

Here, we use Genomic SEM to replicate a recent model of cross-disorder genetic risk and conduct a multivariate GWAS using the largest available GWAS summary statistics of 11 psychiatric disorders. We then compute PRS of the resulting latent genetic factors in the PCA-selected European ancestry subsample of the independent Adolescent Brain Cognitive Development (ABCD) Study sample, from which we conduct a PheWAS testing associations between each PRS with phenotypes measured at the baseline (sample  $n = 5,556$ ; phenotype  $n = 1,269$ ; ages 9-10) and two-year-follow-up (sample  $n = 5,048$ ; phenotype  $n = 1,694$ ; ages 10-13) assessment waves. We conceptualized these phenotypes as being of multiple types: a) more proximal phenotypes that reflect symptoms of various disorders as well as potentially trait-like characteristics that might also reflect psychopathology or related risk factors (e.g., behavioral inhibition); and b) more distal phenotypes that might reflect behavior related to psychopathology or correlated features (e.g., electronic media use, physical health, cognition, etc., **Supplemental Tables S1-S4**). To further mechanistic insight, we also explore whether baseline neural metrics indirectly link PRS to follow-up outcomes (**Supplemental Table S5**). The young age of this

deeply phenotyped sample allows us to investigate early associations with subthreshold symptoms and related traits largely preceding the development of functional impairment, providing etiological insights beyond what is achievable with Electronic Health Records that may prove fruitful for clinical applications that involve early intervention or even prevention.

## RESULTS

See [Figure 1](#) for a summary of the analytic process.

Our multivariate GWAS of 11 psychiatric diagnoses/phenotypes replicated Grotzinger et al.<sup>23</sup> and characterized the genomic architecture of four correlated psychopathology factors ([Fig. 2](#); [Extended Data Figs. 1-3](#); [Supplemental Tables S1-S5](#); [Methods](#)): **1) Compulsive** ( $n_{\text{effective}}=39,364$ ; i.e., OCD, Anorexia, and Tourette Syndrome; 11 genome-wide significant [GWS] loci), **2) Psychotic** ( $n_{\text{effective}}=160,932$ ; i.e., Schizophrenia, Bipolar Disorder, and Problematic Alcohol Use; 224 GWS loci), **3) Neurodevelopmental** ( $n_{\text{effective}}=37,665$ ; i.e., ADHD, Autism Spectrum Disorder, PTSD, MDD, Tourette Syndrome, and Problematic Alcohol Use; 14 GWS loci), and **4) Internalizing** ( $n_{\text{effective}}=432,638$ ; i.e., MDD, GAD, PTSD, and Problematic Alcohol Use; 7 GWS loci).

We conducted a PheWAS of polygenic risk scores (PRS) for these factors (i.e., Compulsive, Psychotic, Neurodevelopmental, and Internalizing PRS) with baseline ( $n_{\text{analytic}}$  up to 5,556; 9.93±0.63 years; 47.0% female) and 2-year follow-up (FU2;  $n_{\text{analytic}}$  up to 5,048; 12.02±0.66 years; 46.93% girls) ABCD data ([Table 1](#)). Bonferroni and False Discovery Rate (FDR) correction were applied to all associations, separately for each PRS (i.e., 1,269 and 1,694 at baseline and FU2, respectively). Analyses revealed multiple associations with Neurodevelopmental PRS ( $n_{\text{fdr}}=637-702$ ;  $n_{\text{Bonferroni}}=187-211$ ; [Fig. 3a](#)) and Internalizing PRS ( $n_{\text{fdr}}=601-687$ ;  $n_{\text{Bonferroni}}=122-173$ ; [Fig. 3b](#)) with few associations identified for Psychotic PRS ( $n_{\text{fdr}}=7$ ;  $n_{\text{Bonferroni}}=1$ ; [Fig. 3d](#)) and none identified for Compulsive PRS ([Fig. 3d](#); [Table 2](#)).

**Neurodevelopmental PRS-PheWAS.** Neurodevelopmental PRS was significantly associated with  $\geq 187$  phenotypes at each timepoint ( $n_{\text{baseline-fdr}}=637$ ;  $n_{\text{baseline-Bonferroni}}=187$ ;  $n_{\text{FU2-fdr}}=702$ ;  $n_{\text{FU2-Bonferroni}}=211$ ; [Table 2](#); [Fig. 3a](#); [Extended Data Figs. 4-5](#), [Supplemental Tables S11-S12](#)). The child mental health and screen time domains were the most frequently associated with Neurodevelopmental PRS across both waves ([Table 2](#)), with proximal (i.e., those mirroring discovery GWAS phenotypes) neurodevelopmental phenotypes including inattention and impulsivity as well as more distal phenotypes such as total screen time, sleep problems, worse relationships with teachers, psychotic-like experiences, neighborhood crime, lower parental monitoring, and receipt of mental health services among the most significantly associated phenotypes. A sizable proportion (13-21%) of cognitive phenotypes were negatively associated with Neurodevelopmental PRS, but notably were not among the top hits and constituted a small overall number of phenotypes.

Of the 799 phenotypes assessed at both waves, 159 ( $n_{\text{FDR}}=442$ ) and 115 ( $n_{\text{FDR}}=366$ ) showed Bonferroni-corrected significant associations at baseline and FU2, respectively, of which 75 ( $n_{\text{FDR}}=284$ ) were consistently associated across waves. Among these phenotypes associated across waves were those reflecting child mental health (e.g., CBCL and teacher-reported Brief

Problem Monitor items and scales reflecting inattentive and externalizing behaviors), special services in school, school performance, psychotic-like experiences summary scores, screen time metrics, sleep issues, and impulsive personality traits (e.g., lack of perseverance, lack of planning). Wave-specific associations (baseline:  $n_{FDR}=349$ ,  $n_{Bonferroni}=112$ ; [Supplemental Table S11](#); FU2:  $n_{FDR}=418$ ,  $n_{Bonferroni}=136$ ; [Supplemental Table S12](#); [Table 2](#)) were predominantly attributable to phenotypes only assessed at either session (e.g., prenatal exposures and family history at baseline; Autism Spectrum symptoms and youth-reported externalizing behaviors at FU2), as well as those associated with broad spectrum risk (e.g., increased youth and caregiver internalizing and broad behavior problems at baseline). As children began to progress into adolescence (FU2), Neurodevelopmental PRS showed more associations with dysfunction that typically become more evident at this age (e.g., worse school performance, conflict at school and with caregiver). Several of these associations (e.g., negative associations with liking school because they do well, feeling as smart as their peers, and school involvement) had non-overlapping 95% confidence intervals across waves ([Supplemental Table S13](#)), though it should be noted that some of these variables (e.g., conflict with caregiver about grades) are more frequently endorsed at FU2.

**Internalizing PRS-PheWAS.** Polygenic risk for the Internalizing factor was significantly associated with  $\geq 122$  phenotypes at each timepoint ( $n_{baseline-fdr}=601$ ;  $n_{baseline-Bonferroni}=122$ ;  $n_{FU2-fdr}=687$ ;  $n_{FU2-Bonferroni}=173$ ; [Table 2](#); [Fig. 3b](#); [Extended Data Figs. 4-5](#), [Supplemental Tables S11-S12](#)). The child mental health, family mental health, and screen time domains were most frequently associated with Internalizing PRS, with Bonferroni-corrected significant associations ranging from 8.4% (screen time at FU2) to 22% (screen time at baseline) of assessed phenotypes. The other domains showed less frequent associations, with between 0% (culture/environment at baseline, and substance-related and cognition phenotypes at FU2) and 7.7% (Demographics at FU2) of associations surviving Bonferroni correction ([Table 2](#)). Top associations included proximal internalizing phenotypes, such as depressive and stress-related symptoms, as well as more distal phenotypes including total behavior problems, psychotic-like experiences, sleep difficulties, and externalizing behaviors.

Of the 799 phenotypes assessed at both waves, 98 ( $n_{FDR}=414$ ) and 100 ( $n_{FDR}=386$ ) were significantly (Bonferroni) associated with Internalizing PRS at baseline and FU2, respectively, of which 46 ( $n_{FDR}=275$ ) were consistently associated across both waves. Featured among these phenotypes were those reflecting internalizing, externalizing, and total behavior problems; psychotic-like experiences, sleep disturbances, and school difficulty and special services. Associations specific to each wave (baseline:  $n_{FDR}=326$ ,  $n_{Bonferroni}=76$ ; [Supplemental Table S11](#); FU2:  $n_{FDR}=412$ ,  $n_{Bonferroni}=172$ ; [Supplemental Table S12](#); [Table 2](#)) largely reflected phenotypes assessed only at one timepoint (e.g., family history and prenatal exposures at baseline; youth-reported behavior problems, items on the Early Adolescent Temperament Questionnaire, stressful life events at FU2), as well as additional school-related risk factors (e.g., drop in grades, detention/suspension, self-reported dislike of school and feeling less smart than peers) and caregiver-reported mood swings at FU2. As with Neurodevelopmental PRS, 95% confidence intervals for several of these associations were non-overlapping across waves, with stronger associations observed at FU2 including both feeling less smart than peers and psychotic-like experience items, which may overlap in content with social anxiety characteristics (i.e.,

suddenly not trusting others who seem to be watching or talking about them, difficulty figuring out how to say something so others can understand; [Supplemental Table S13](#)).

**Compulsive PRS-PheWAS.** Polygenic risk for the Compulsive factor was not significantly associated with any phenotypes at either wave following FDR and Bonferroni correction ([Fig. 3c](#); [Extended Data Figs. 4-5](#), [Supplemental Tables S11-S12](#)). Nominally significant associations (i.e.,  $p < 0.05$ ) included past OCD diagnosis and symptoms at baseline as well as negative associations with caregiver self-reported poor diet, time spent watching TV on the weekend, and measured weight and waist circumference at FU2.

**Psychotic PRS-PheWAS.** Polygenic risk for the Psychotic factor was significantly associated with lower school involvement at FU2 ( $n_{\text{baseline-fdr}}=0$ ;  $n_{\text{baseline-Bonferroni}}=0$ ;  $n_{\text{FU2-fdr}}=7$ ;  $n_{\text{FU2-Bonferroni}}=1$ ; [Fig. 3d](#); [Extended Data Figs. 4-5](#), [Supplemental Tables S11-S12](#)). Nominally significant associations included functional impairment due to compulsions, caregiver history with mania, and negative associations with child-reported school factors (e.g., feeling as smart as peers, liking school because they are doing well). [Figure 4](#) summarizes primary PheWAS results.

**Imaging Analyses.** Compulsive PRS was significantly and positively associated with three imaging metrics: left precentral volume and surface area, and right transverse temporal gyrus ( $\beta_s \geq 0.032$ ,  $p_s \leq 5.6e-04$ ). Psychotic PRS showed significant negative associations with right pars orbitalis thickness, right ventral diencephalon volume, and fractional anisotropy of the corpus callosum ( $\beta_s \leq -0.017$ ,  $p_s \leq 1.1e-03$ ). Neurodevelopmental PRS was broadly negatively associated with global volume metrics (i.e., total, total right, and total left hemisphere cortical volume, whole brain volume, supratentorial volume), right mean thickness, and 4<sup>th</sup> ventricle volume ( $\beta_s \leq -0.037$ ,  $p_s \leq 2.6e-03$ ). Internalizing PRS was not associated with any imaging metrics following Bonferroni correction ([Supplemental Tables S14-S21](#); [Extended Data Figs. 6-10](#)).

**Indirect Effects Analyses.** To explore whether brain metrics may indirectly link PRS to two-year follow-up phenotypes, associations between each of the significantly associated imaging measures and each of the significantly associated phenotypes at follow-up were examined (i.e., Psychotic PRS: 3 imaging metrics \* 1 phenotype; Neurodevelopmental PRS: 5 imaging metrics \* 211 phenotypes). None of the imaging metrics associated with Psychotic PRS were significantly linked to school involvement, the single phenotype whose association with Psychotic PRS survived Bonferroni correction. For Neurodevelopmental PRS, 47-51 FU2 phenotypes showed significant associations with total cortical volume, supratentorial volume, and whole brain volume ( $|\beta_s| \leq 0.07$ ,  $p_s \leq 4.5e-04$ ) following Bonferroni correction within each imaging metric. The strongest associations were between all three volume indices and indicators of school performance, with smaller volumes being linked to worse performance. Other phenotypes negatively associated with brain volumes included CBCL externalizing symptoms, issues with concentration, symptoms of mania, total weekday screen time, and being on Medicaid insurance. Path analysis models including the direct effect of PRS on follow-up phenotype in addition to the indirect effect of PRS on follow-up phenotype via cortical volume were fit. Only one association survived Bonferroni correction (for 47-51 tests, depending on metric): Total cortical volume indirectly linked Neurodevelopmental PRS to grades in school. Neurodevelopmental PRS was negatively associated with total cortical volume ( $\beta = -0.056$ , 95%

CI [-0.087, -0.026]), which was negatively associated with worse grades in school ( $\beta=-0.14$ , 95% CI [-0.18, -0.11]), and indirectly linked Neurodevelopmental PRS to worse grades in school (direct effect  $\beta=0.012$ ; indirect effect  $\beta=0.0081$ ,  $p=8.9e-04$ ; **Figure 5**). Although the majority of other indirect effects were significant at the 0.05 level, none survived correction for the total number of tests. All results are represented in **Supplemental Tables S22-S23**.

## DISCUSSION

Among 5,048-5,556 youth of European ancestry aged 9-13, PRS for broad spectrum Neurodevelopmental and Internalizing psychopathology were associated with numerous proximal phenotypes (e.g., inattention and impulsivity; depressive symptoms and anxiety) as well as a host of more distal variables (e.g., metrics of screen time; stressful life events). Compulsive and Psychotic PRS largely failed to predict even proximal phenotypes, with Psychotic PRS showing only one significant association with lower school involvement. Several significant associations between PRS and neuroimaging metrics emerged, such as between Neurodevelopmental PRS and global volume measures. Lower total cortical volume measured at baseline partially indirectly linked Neurodevelopmental PRS to lower grades in school assessed at two-year follow-up. This study builds upon prior PRS-PheWAS by leveraging PRS of higher order latent factors representing phenotypes across the psychiatric diagnostic spectrum, an approach that enhances the power of PRS to explain meaningful variation in proximal phenotypes and better captures the highly polygenic nature of psychopathology risk compared to individual disorder PRS.

A first key finding from this work is that both Neurodevelopmental and Internalizing PRS were associated with a range of neurodevelopmental and internalizing phenotypes across waves, far more so than Compulsive and Psychotic PRS. In particular, across both waves, Neurodevelopmental PRS was consistently and strongly associated with behaviors and symptoms of inattention, impulsivity, and externalizing disorders as reported by caregivers, teachers, and the youth themselves. For Internalizing PRS, top hits included depressive, internalizing, and stress-related symptoms, as well as family history of depression and anxiety. These results provide validation to the multivariate GWAS PRS-PheWAS approach to detecting and replicating associations with proximal phenotypes.

Neurodevelopmental and Internalizing PRS exhibited both dissociable associations and substantial convergence, with 33%-57% of significant hits at both waves shared across these PRS. Top hits with Internalizing PRS were broader in nature, including associations with general and externalizing psychopathology symptoms (e.g., on the CBCL and Brief Problem Monitor measures), psychotic-like experiences, sleep disorder issues, and stressful life events of similar magnitude to associations with internalizing symptoms. Although Neurodevelopmental PRS was linked to multiple phenotypes across these domains, top hits were more constrained to Neurodevelopmental disorder symptoms and behaviors related to impulsivity and impaired attention. This pattern of results converges with phenotypic findings in the ABCD sample that stressful life events are more strongly related to internalizing symptoms, whereas contextual variables related to externalizing behaviors, such as school factors, are more strongly related to externalizing symptoms<sup>13</sup>. At the same time, the substantial overlap in associations aligns with

previous findings in this sample showing that latent neurodevelopmental and internalizing factors are moderately correlated with one another, and, further, that general and domain-specific factors together explain the most variance in psychopathology symptoms, school performance, and measures of cognitive ability<sup>14</sup>. Overall, the broad convergence across genetic liability for psychopathology spectra not only includes psychopathology phenotypes but also phenotypes overlapping with cognition and physical health.

Beyond demonstrating that many previously identified phenotypic associations hold at the level of polygenic risk for higher order psychopathology factors in children prior to the typical expression of psychiatric disorder, the present study further highlights associations with more distal phenotypes that may pave the way for future etiologic understanding. Several substance-related phenotypes were associated with Neurodevelopmental PRS, including prenatal tobacco exposure, ease of obtaining cigarettes and marijuana, and peer approval of use of cigarettes and marijuana. These findings add to prior work showing that ADHD PRS is linked to tobacco use disorder in adults<sup>15</sup> and that “deviant” peer affiliation may link ADHD to future substance use<sup>16</sup>. These associations may be in part due potential gene-environment correlation between parental liability to Neurodevelopmental disorders and raising children in environments with increased access to substances. The large number of associations between Neurodevelopmental PRS and metrics of screen time replicates a recent finding linking ADHD PRS to total screen time use in the ABCD sample<sup>11</sup> and offers further insight into specific screen media behaviors. In the present analyses, there was evidence consistent with the idea that genetic vulnerability for neurodevelopmental concerns may increase susceptibility to the rewarding aspects of screen media (e.g., screen time as a source of motivation) as well as increase aspects of problematic use (e.g., struggling to limit screen time).

The numerous associations between Internalizing PRS and sleep difficulties, psychotic-like experiences, and stressful life events and peer victimization also highlight fruitful avenues for further exploration. Diagnostic criteria for depression and anxiety include sleep difficulties, which, according to the present analyses, may be particularly salient and modifiable risk factors for internalizing disorders. Psychotic-like experiences have previously been found to show strong associations with broad psychopathology in the ABCD study sample, including internalizing symptoms, and the present findings strengthen the notion that psychotic-like experiences may constitute a broad manifestation of psychopathology risk in childhood<sup>17</sup>. Furthermore, associations between internalizing symptom genetic liability and stressful life events can be interpreted as being consistent with gene-environment correlation and stress generation theories and evidence that parental depression may contribute to stress experienced by children<sup>18,19</sup>. Associations between Internalizing PRS and self-reported peer victimization phenotypes (e.g., peer gossip and threats to reputation) are also consistent with the theory that relational stressors may be experienced relatively more frequently by those at risk for internalizing symptoms<sup>18</sup>.

By contrast, PRS for Compulsive and Psychotic disorders showed zero and one significant associations, respectively. There are several non-mutually exclusive explanations for this pattern of findings. *First*, the psychiatric phenotypes comprising the Compulsive and Psychotic factors (e.g., OCD, Anorexia, Schizophrenia, Bipolar Disorder) are generally less

prevalent than disorders such as ADHD, Generalized Anxiety, and Major Depression<sup>1,20</sup>. As such, the ABCD sample, which was recruited to be representative of the United States 9–10-year-old population, may be insufficiently enriched for Compulsive and Psychotic disorders for associations to be detected with genetic liability. *Second*, and relatedly, disorders such as Schizophrenia and Bipolar Disorder in particular, and to a lesser degree OCD and Eating Disorders, typically manifest later than ADHD and anxiety. It is possible that associations with Compulsive and Psychotic PRS may arise as the ABCD sample ages. Notably, the relative lack of findings for Psychotic PRS mirrors prior work in the ABCD sample that found that Schizophrenia PRS were not significantly associated with continuously measured psychotic-like experiences, whereas cross-disorder PRS were<sup>21</sup>. It is possible that phenotypic expression of underlying genetic liability to these domains may be observed as the sample ages, which would converge with a recent PheWAS documenting associations between PRS for Schizophrenia and psychotic symptoms in adults<sup>15</sup>.

A second set of analyses examined associations between each PRS and brain structure and resting state functional connectivity, adding to the emerging neurogenetics literature. Findings for Compulsive PRS include a positive association with right transverse temporal cortex surface area, notable in its direction given prior evidence in the ENIGMA OCD sample (and lack of association in the ENIGMA study of children and adolescent OCD), with the present sample perhaps pointing to the influence of maturational changes specific to middle childhood.<sup>22</sup> OCD has been previously linked to increased intra-frontoparietal network connectivity<sup>23,24</sup>, as in the present study, and may reflect over control and excessive monitoring tendencies typical of compulsive disorders. Psychotic PRS was associated with lower pars orbitalis thickness, lower ventral diencephalon volume, and lower fractional anisotropy of the corpus callosum, consistent with previous research on clinical high risk for psychosis (though pars orbitalis thickness may be associated especially with more chronic disease stages)<sup>25–27</sup>. Neurodevelopmental PRS was negatively associated with global gray matter volume, as well as right cortical thickness, whereas Internalizing PRS showed no significant associations with neural metrics. This pattern aligns with recent work showing widespread negative associations between externalizing, but not internalizing, symptoms and cortical and subcortical regional gray matter volume, although notably the majority of significant associations did not survive covarying for total intracranial volume<sup>28</sup>.

Given emerging evidence that brain structure may mechanistically link genetic liability to phenotypic expression<sup>e.g.,9,29</sup>, we probed whether baseline neural metrics assessed significantly associated with PRS indirectly link said PRS with significantly associated phenotypes assessed at two-year follow-up. We found that smaller total cortical volume indirectly linked Neurodevelopmental PRS to poorer grades in school, extending prior work showing that brain volume is positively associated with educational performance<sup>30</sup>. Global gray matter structural alterations may be one mechanism through which polygenic liability to neurodevelopmental disorders may impact school performance, and although future analyses should explore putative environmental and other mechanisms. It should also be noted that generally small genetic correlations between psychopathology and neural metrics may constrain power to detect these associations<sup>31</sup>.

## Limitations

The major limitation of the present study is the exclusion of genetic ancestry groups other than European from analyses. Although this constraint is attributable to the lack of available GWAS summary statistics on the 11 psychiatric phenotypes comprising the multivariate GWAS in other genetic ancestry groups and the poor predictive ability of PRS across genetic ancestry<sup>32</sup>, the applicability of these findings only to those of European genetic ancestry has the potential to increase disparities in healthcare interventions and access to information based on ancestry. All findings described above should be interpreted with caution, as they may have poor generalizability beyond Americans of European genetic ancestry.

In addition, these latent higher order genetic psychopathology factors may be missing important diagnoses, such as social anxiety, conduct, and oppositional defiant disorders, as GWAS on these disorders are either not available or are severely underpowered. The Compulsive factor consisted of the three disorders with lowest  $N_s$ , which may have weakened power for PRS development and discovery. It is worth noting, however, that power is not the only driver of discovery, as three of the largest samples loaded on the psychotic factor, which arguably had more power to train PRS but was significantly associated with fewer phenotypes than the Neurodevelopmental or Internalizing PRS. Numerous associations were observed in the context of a relatively small Neurodevelopment GWAS effective sample size ( $N_{\text{Effective}}=37,665$ ), perhaps reflecting the developmentally normative nature of neurodevelopmental phenotypes observed in this age group. Further, although we elected to use strict p-value correction (i.e., Bonferroni) given the large number of tests and exploratory nature of analyses, this method may be overly stringent in ignoring the large pattern of correlation among phenotypes, and therefore we also provide FDR correction estimates in all results tables. Finally, although the present analyses do not extend into peak ages of psychiatric risk for most disorders, it is important to identify phenotypes in childhood that may represent critical signatures of genetic risk for psychopathology.

## Conclusions

This study builds on prior psychiatric PRS-PheWAS<sup>15,33,34</sup> by interrogating associations between PRS derived from higher order latent factors representing phenotypes across the diagnostic spectrum in a sample of youth. The young age of the sample and related low levels of diagnosable psychiatric disorders enabled the identification of putative risk factors associated with genetic liability largely prior to disorder onset. Further, in contrast to the Electronic Health Records approach of many PheWAS studies, the present study leveraged deeply characterized phenotypes encompassing both threshold and subthreshold psychopathology symptoms as well as a host of other domains not typically included in EHR data (e.g., brain structure, screen time, environmental exposures, multiple reporters and modes of assessment). The present study provides support for the power of agnostic methods to improve understanding of the etiopathogenesis of psychopathology in youth.



### **Acknowledgements:**

SEP, NRK, DMB, and RB developed the research questions. SEP, AJG, NRK, IN, LB, ISH, and NME cleaned the phenotypic data. SEP, SC, AJG, and NRK conducted analyses. SEP and NRK drafted the manuscript. SC, AJG, ISH, ECJ, ASH, DAAB, NME, DB, and RB provided critical revision of the manuscript for important intellectual content. SEP and NRK had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

### **Code availability:**

Code for Genomic SEM and PheWAS analyses are located on GitHub: <https://github.com/sepaul1124/ABCD-PHEWAS>

### **Funding:**

Data for this study were provided by the Adolescent Brain Cognitive Development (ABCD) study<sup>□</sup>, which was funded by awards U01DA041022, U01DA041025, U01DA041028, U01DA041048, U01DA041089, U01DA041093, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147 from the NIH and additional federal partners (<https://abcdstudy.org/federal-partners.html>). This study was supported by R01 DA054750. Authors received funding support from NIH: SEP was supported by F31AA029934. NRK was supported by K23MH12179201. AJG was supported by NSF DGE-213989. ECJ was supported by K01DA051759. ASH was supported by K01AA030083. DMB (R01-MH113883; R01-MH066031; U01-MH109589; U01-A005020803; R01-MH090786), RB (R01-DA054750, R01-AG045231, R01-AG061162, R21-AA027827, R01-DA046224, U01-DA055367). NME was supported by NSF DGE-1745038. The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

### **Conflicts of interest:**

The authors do not report any conflicts of interest.

**Table 1.** Sample characteristics.

<b>Variable</b>	<b>Mean (SD) / n (%)</b>
Sex (female)	2,612 (47.0%)
Age (years)	9.93 (0.63)
<b>Household Income</b>	
< \$35,000	375 (7.1%)
\$35,000 - \$49,999	283 (5.3%)
\$50,000 - \$74,999	717 (13.5%)
\$75,000 - \$99,999	896 (16.9%)
\$100,000 - \$199,999	2,178 (41.1%)
≥ \$200,000	849 (16.0%)
<b>Highest Caregiver Education</b>	
Less than high school	25 (0.45%)
High school degree or equivalent	188 (3.4%)
Some college, associate degree	1046 (18.8%)
College degree	1,753 (31.6%)
Master's degree	1,723(31.0%)
Doctorate/professional degree	820 (14.8%)
<b>Caregiver Marital Status</b>	
Married	4,566 (82.2%)
Divorced	490 (8.8%)
Widowed	41 (0.74%)
Separated	131 (2.4%)
Never married	164 (3.0%)
Living with partner	161 (2.9%)
<b>Child Behavior Checklist (CBCL) T Scores</b>	
Internalizing	48.7 (10.4)
Externalizing	45.4 (10.0)
Total problems	45.8 (10.8)

*Note.* Measures included here are included to describe the sample at baseline. Household income, caregiver education, and marital status were not included in analyses. CBCL raw scores, rather than T scores, were used in analyses.

**Table 2.** Summary of significant associations.

		<b>Baseline</b>							
		Compulsive		Psychotic		Neurodevel.		Internalizing	
	Total	FDR	Bonf	FDR	Bonf	FDR	Bonf	FDR	Bonf
Child MH	640	0 (0%)	0 (0%)	0 (0%)	0 (0%)	403 (63%)	144 (23%)	354 (55%)	74 (12%)
Fam MH	239	0 (0%)	0 (0%)	0 (0%)	0 (0%)	97 (41%)	11 (4.6%)	128 (54%)	26 (11%)
Physical Health	172	0 (0%)	0 (0%)	0 (0%)	0 (0%)	53 (31%)	9 (5.2%)	59 (34%)	13 (7.6%)
Environment	113	0 (0%)	0 (0%)	0 (0%)	0 (0%)	36 (32%)	3 (2.7%)	17 (15%)	0 (0%)
Demographics	27	0 (0%)	0 (0%)	0 (0%)	0 (0%)	10 (37%)	3 (11%)	10 (37%)	1 (3.7%)
Substance	46	0 (0%)	0 (0%)	0 (0%)	0 (0%)	15 (33%)	4 (8.7%)	11 (24%)	3 (6.5%)
Screen Time	18	0 (0%)	0 (0%)	0 (0%)	0 (0%)	13 (72%)	10 (56%)	14 (78%)	4 (22%)
Cognition	14	0 (0%)	0 (0%)	0 (0%)	0 (0%)	10 (71%)	3 (21%)	9 (64%)	1 (7.1%)

		<b>Two-Year Follow-Up</b>							
		Compulsive		Psychotic		Neurodevel.		Internalizing	
	Total	FDR	Bonf	FDR	Bonf	FDR	Bonf	FDR	Bonf
Child MH	636	0 (0%)	0 (0%)	2 (0.3%)	0 (0%)	300 (47%)	134 (21%)	315 (50%)	108 (17%)
Fam MH	158	0 (0%)	0 (0%)	1 (0.6%)	0 (0%)	45 (28%)	0 (0%)	98 (62%)	19 (12%)
Physical Health	199	0 (0%)	0 (0%)	0 (0%)	0 (0%)	67 (34%)	11 (5.5%)	66 (33%)	9 (4.5%)
Environment	441	0 (0%)	0 (0%)	3 (0.7%)	1 (0.2%)	162 (37%)	39 (8.8%)	125 (28%)	26 (5.9%)
Demographics	26	0 (0%)	0 (0%)	0 (0%)	0 (0%)	11 (42%)	2 (7.7%)	7 (27%)	2 (7.7%)
Substance	119	0 (0%)	0 (0%)	0 (0%)	0 (0%)	53 (45%)	5 (4.2%)	20 (17%)	0 (0%)
Screen Time	107	0 (0%)	0 (0%)	1 (0.9%)	0 (0%)	60 (56%)	19 (18%)	54 (50%)	9 (8.4%)
Cognition	8	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (63%)	1 (13%)	2 (25%)	0 (0%)

*Note.* Summary of the number of significant associations per domain, at each wave of assessment. The total number of included phenotypes per domain are represented in the second column. The number and percent (of the total number of phenotypes in that domain) of significant associations (FDR- and Bonferroni-corrected p-values) are depicted. Bonf. = Bonferroni. Neurodevel. = Neurodevelopmental. Percentages are rounded to two digits.

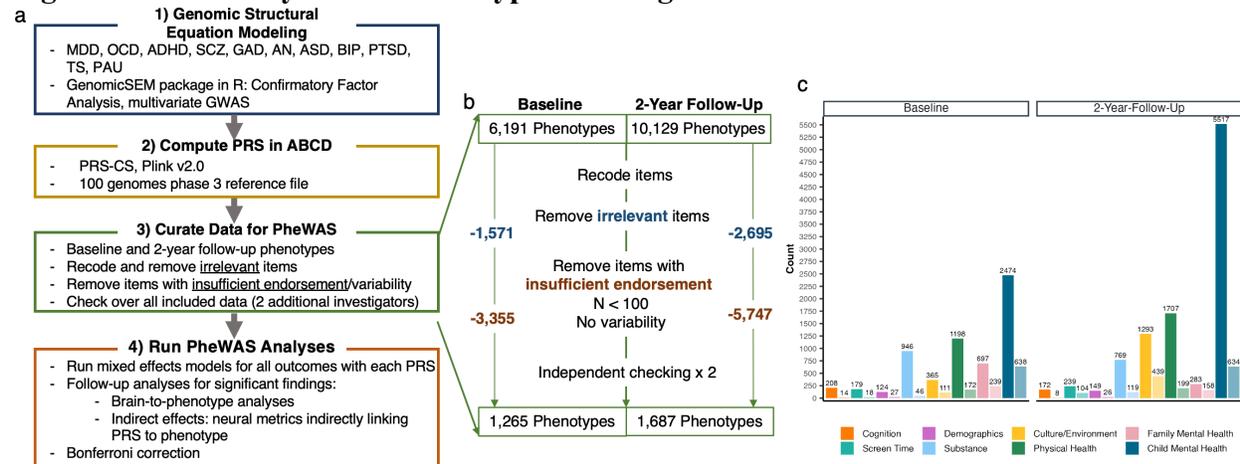
## References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593-602. doi:10.1001/archpsyc.62.6.593
2. Kyu HH, Abate D, Abate KH, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018;392(10159):1859-1922. doi:10.1016/S0140-6736(18)32335-3
3. Pettersson E, Lichtenstein P, Larsson H, et al. Genetic influences on eight psychiatric disorders based on family data of 4 408 646 full and half-siblings, and genetic data of 333 748 cases and controls. *Psychol Med*. 2019;49(7):1166-1173. doi:10.1017/S0033291718002039
4. Wendt FR, Pathak GA, Tylee DS, Goswami A, Polimanti R. Heterogeneity and Polygenicity in Psychiatric Disorders: A Genome-Wide Perspective. *Chronic Stress*. 2020;4:2470547020924844. doi:10.1177/2470547020924844
5. Kotov R, Krueger RF, Watson D, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): A Quantitative Nosology Based on Consensus of Evidence. *Annu Rev Clin Psychol*. 2021;17:83-108. doi:10.1146/annurev-clinpsy-081219-093304
6. Cuthbert BN. Research Domain Criteria (RDoC): Progress and Potential. *Curr Dir Psychol Sci*. 2022;31(2):107-114. doi:10.1177/096372142111051363
7. Grotzinger AD, Rhemtulla M, Vlaming R. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav*. 2019;3(5):513-525. doi:10.1038/s41562-019-0566-x
8. Kotov R, Krueger RF, Watson D, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J Abnorm Psychol*. 2017;126(4):454-477. doi:10.1037/abn0000258
9. Lett TA, Vogel BO, Ripke S, et al. Cortical Surfaces Mediate the Relationship Between Polygenic Scores for Intelligence and General Intelligence. *Cereb Cortex N Y NY*. 2020;30(4):2708-2719. doi:10.1093/cercor/bhz270
10. Sudre G, Frederick J, Sharp W, et al. Mapping associations between polygenic risks for childhood neuropsychiatric disorders, symptoms of attention deficit hyperactivity disorder, cognition, and the brain. *Mol Psychiatry*. 2020;25(10):2482-2492. doi:10.1038/s41380-019-0350-3
11. Yang A, Rolls ET, Dong G, et al. Longer screen time utilization is associated with the polygenic risk for Attention-deficit/hyperactivity disorder with mediation by brain white matter microstructure. *eBioMedicine*. 2022;80. doi:10.1016/j.ebiom.2022.104039

12. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene–disease associations. *Bioinformatics*. 2010;26(9):1205-1210. doi:10.1093/bioinformatics/btq126
13. Choi KW, Wilson M, Ge T, et al. Integrative analysis of genomic and exposomic influences on youth mental health. *J Child Psychol Psychiatry*. 2022;63(10):1196-1205. doi:10.1111/jcpp.13664
14. Michelini G, Barch DM, Tian Y, Watson D, Klein DN, Kotov R. *Delineating and Validating Higher-Order Dimensions of Psychopathology in the Adolescent Brain Cognitive Development (ABCD) Study*. *Epidemiology*; 2019. doi:10.1101/548057
15. Kember RL, Merikangas AK, Verma SS, et al. Polygenic Risk of Psychiatric Disorders Exhibits Cross-trait Associations in Electronic Health Record Data From European Ancestry Individuals. *Biol Psychiatry*. 2021;89(3):236-245. doi:10.1016/j.biopsych.2020.06.026
16. Marshal MP, Molina BSG, Pelham WE. Childhood ADHD and Adolescent Substance Use: An Examination of Deviant Peer Group Affiliation as a Risk Factor. *Psychol Addict Behav J Soc Psychol Addict Behav*. 2003;17(4):293-302. doi:10.1037/0893-164X.17.4.293
17. Karcher NR, Loewy RL, Savill M, et al. Replication of Associations With Psychotic-Like Experiences in Middle Childhood From the Adolescent Brain Cognitive Development (ABCD) Study. *Schizophr Bull Open*. 2020;1(1):sgaa009. doi:10.1093/schizbullopen/sgaa009
18. Adrian C, Hammen C. Stress exposure and stress generation in children of depressed mothers. *J Consult Clin Psychol*. 1993;61:354-359. doi:10.1037/0022-006X.61.2.354
19. Kendler KS, Karkowski LM, Prescott CA. Causal Relationship Between Stressful Life Events and the Onset of Major Depression. *Am J Psychiatry*. 1999;156(6):837-841. doi:10.1176/ajp.156.6.837
20. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*. 2022;9(2):137-150. doi:10.1016/S2215-0366(21)00395-3
21. Karcher NR, Paul SE, Johnson EC, et al. Psychotic-like Experiences and Polygenic Liability in the Adolescent Brain Cognitive Development Study. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2022;7(1):45-55. doi:10.1016/j.bpsc.2021.06.012
22. Boedhoe PSW, Schmaal L, Abe Y, et al. Cortical Abnormalities Associated With Pediatric and Adult Obsessive-Compulsive Disorder: Findings From the ENIGMA Obsessive-Compulsive Disorder Working Group. *Am J Psychiatry*. 2018;175(5):453-462. doi:10.1176/appi.ajp.2017.17050485
23. Göttlich M, Krämer UM, Kordon A, Hohagen F, Zurowski B. Decreased limbic and increased fronto-parietal connectivity in unmedicated patients with obsessive-compulsive disorder. *Hum Brain Mapp*. 2014;35(11):5617-5632. doi:10.1002/hbm.22574

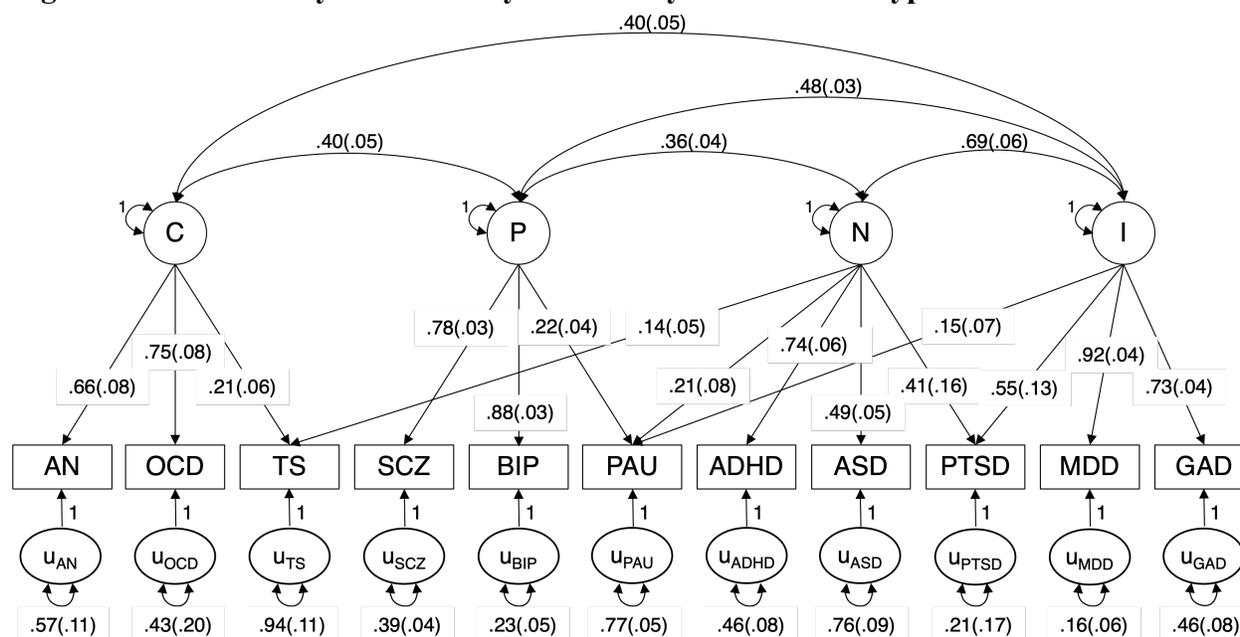
24. Gürsel DA, Reinholz L, Bremer B, et al. Frontoparietal and salience network alterations in obsessive–compulsive disorder: insights from independent component and sliding time window analyses. *J Psychiatry Neurosci JPN*. 2020;45(3):214. doi:10.1503/jpn.190038
25. Remiszewski N, Bryant JE, Rutherford SE, et al. Contrasting Case-Control and Normative Reference Approaches to Capture Clinically Relevant Structural Brain Abnormalities in Patients With First-Episode Psychosis Who Are Antipsychotic Naive. *JAMA Psychiatry*. 2022;79(11):1133-1138. doi:10.1001/jamapsychiatry.2022.3010
26. Smigielski L, Stämpfli P, Wotruba D, et al. White matter microstructure and the clinical risk for psychosis: A diffusion tensor imaging study of individuals with basic symptoms and at ultra-high risk. *NeuroImage Clin*. 2022;35:103067. doi:10.1016/j.nicl.2022.103067
27. ENIGMA Clinical High Risk for Psychosis Working Group. Association of Structural Magnetic Resonance Imaging Measures With Psychosis Onset in Individuals at Clinical High Risk for Developing Psychosis: An ENIGMA Working Group Mega-analysis. *JAMA Psychiatry*. 2021;78(7):753-766. doi:10.1001/jamapsychiatry.2021.0638
28. Blok E, Geenjaar EPT, Geenjaar EAW, Calhoun VD, White T. Neurodevelopmental Trajectories in Children With Internalizing, Externalizing and Emotion Dysregulation Symptoms. *Front Psychiatry*. 2022;13. Accessed December 9, 2022. <https://www.frontiersin.org/articles/10.3389/fpsy.2022.846201>
29. Barker ED, Ing A, Biondo F, et al. Do ADHD-impulsivity and BMI have shared polygenic and neural correlates? *Mol Psychiatry*. 2021;26(3):1019. doi:10.1038/s41380-019-0444-y
30. Ma Q, Wang H, Rolls ET, et al. Lower gestational age is associated with lower cortical volume and cognitive and educational performance in adolescence. *BMC Med*. 2022;20(1):424. doi:10.1186/s12916-022-02627-3
31. Grasby KL, Jahanshad N, Painter JN. The genetic architecture of the human cerebral cortex. *Science*. 2020;367(6484). doi:10.1126/science.aay6690
32. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x
33. Leppert B, Millard LAC, Riglin L, et al. A cross-disorder PRS-pheWAS of 5 major psychiatric disorders in UK Biobank. *PLoS Genet*. 2020;16(5):e1008185. doi:10.1371/journal.pgen.1008185
34. Shen X, Howard DM, Adams MJ, et al. A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. *Nat Commun*. 2020;11(1):2301. doi:10.1038/s41467-020-16022-0

## Figure 1. Data Analysis and Phenotype Curating Flow Chart



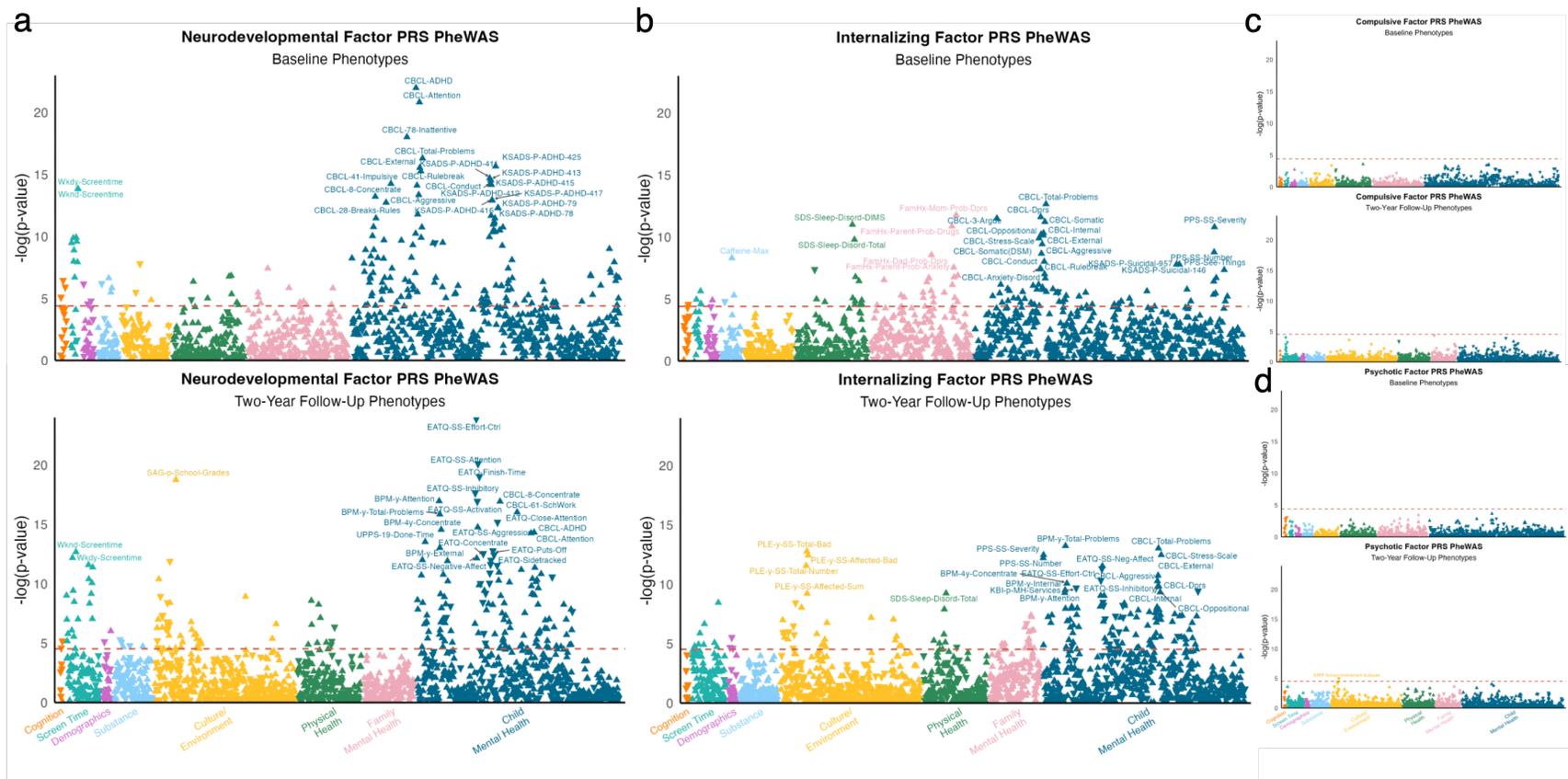
**a.** Flow chart depicting data analytic process. **b.** Flow chart depicting curating of phenotypes in the Adolescent Brain Cognitive Development (ABCD) dataset. **c.** Bar plot illustrating the total number of phenotypes per domain considered for analysis and retained for the final PheWAS. The left and right panels depict phenotypes at baseline and two-year follow-up assessment waves, respectively. The y-axis reflects the number of phenotypes, and the x-axis shows the domains of phenotypes considered. The solid, larger bars depict the total number of phenotypes considered for that domain, whereas the more transparent, shorter bars depict the total number of phenotypes retained for analyses.

**Figure 2. Confirmatory Factor Analysis of 11 Psychiatric Phenotypes.**



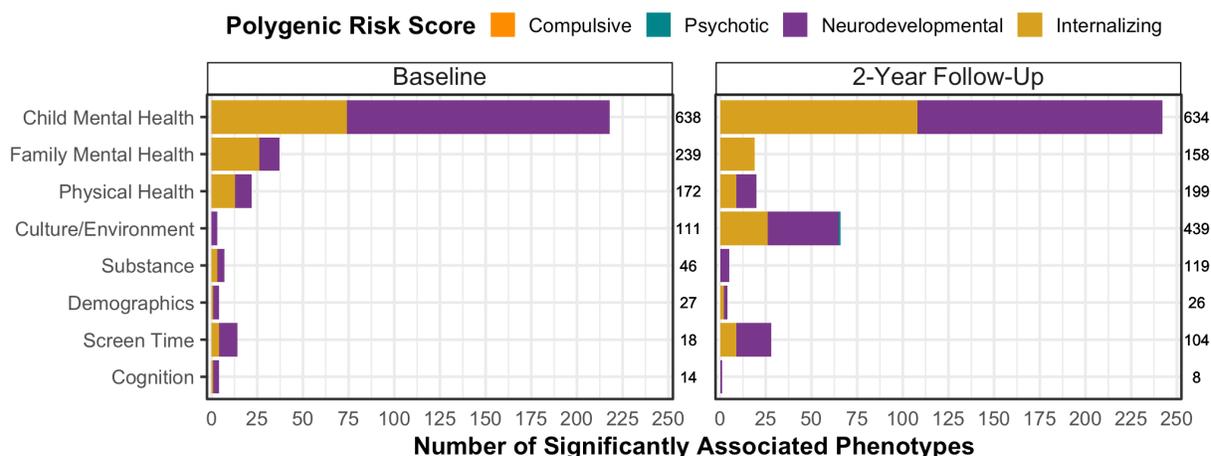
Confirmatory factor analysis model. Loadings and covariances are standardized. Psychiatric phenotype manifest indicators are depicted in rectangles; inferred common genetic components are represented by circles. AN = Anorexia Nervosa; OCD = Obsessive-Compulsive Disorder; TS = Tourette Syndrome; SCZ = Schizophrenia; BIP = Bipolar Disorder; PAU = Problematic Alcohol Use; ADHD = Attention-Deficit/Hyperactivity Disorder; ASD = Autism Spectrum Disorder; PTSD = Post-Traumatic Stress Disorder; MDD = Major Depressive Disorder; GAD = Generalized Anxiety Disorder.

**Figure 3. Manhattan plots of PheWAS results...**



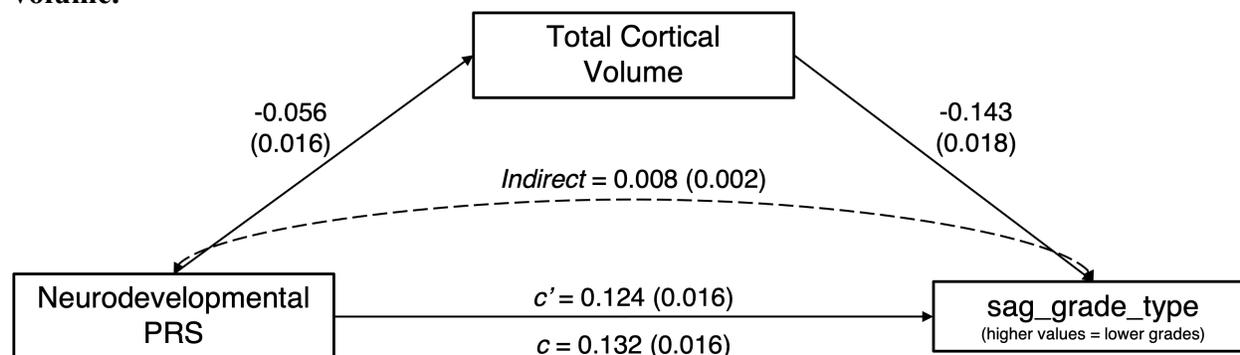
**a.** Neurodevelopmental PRS-PheWAS results (top: Baseline; bottom: FU2). **b.** Internalizing PRS-PheWAS results (top: Baseline; bottom: FU2). **c.** Compulsive PRS-PheWAS results (top: Baseline; bottom: FU2). **d.** Psychotic PRS-PheWAS results (top: Baseline; bottom: FU2). The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with the colors indicating categories and triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. Full results can be found in [Supplemental Tables S11-S12](#).

**Figure 4. Summary of PheWAS findings.**



Summary of phenotypes significantly associated with each PRS, grouped by phenotype domain. The left y-axis reflects the phenotype domains. The right y-axes show the total number of phenotypes analyzed per domain. The x-axis denotes the number of phenotypes significantly associated with each PRS.

**Figure 5. Indirect effect of Neurodevelopmental PRS on school grades through cortical volume.**



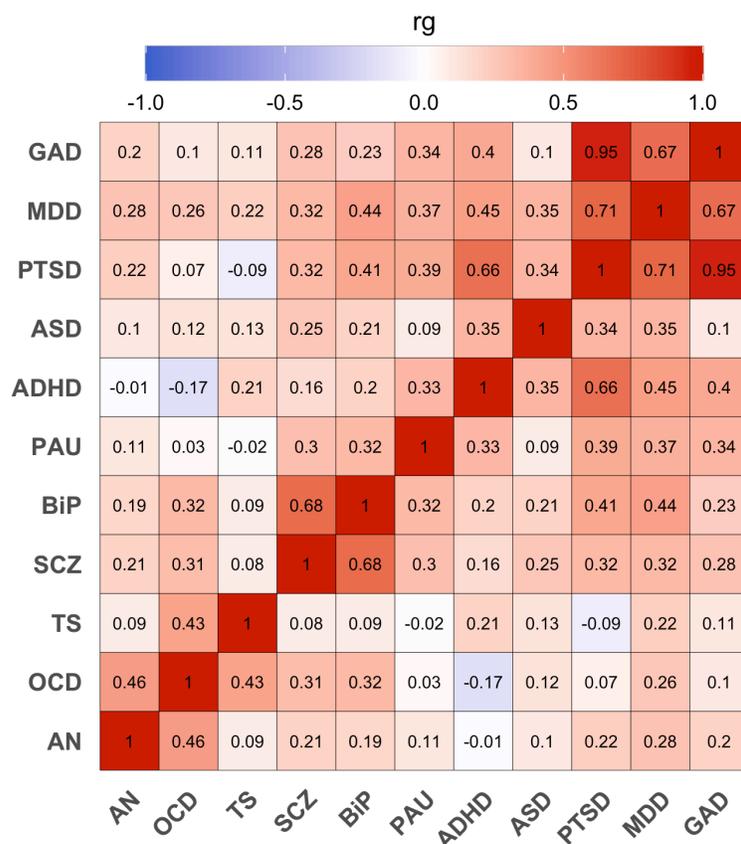
All estimates are standardized and statistically significant ( $p_{BONFERRONI} < 0.05$ ).  $c$  = total effect;  $c'$  = direct effect.

**Table 1.** Sample characteristics.

<b>Variable</b>	<b>Mean (SD) / n (%)</b>
Sex (female)	2,176 (46.8%)
Age (years)	9.9 (0.63)
Household Income	
< \$35,000	963 (21.6%)
\$35,000 - \$49,999	247 (5.5%)
\$50,000 - \$74,999	628 (14.1%)
\$75,000 - \$99,999	791 (17.8%)
\$100,000 - \$199,999	1,826 (41.0%)
Highest Caregiver Education	
Less than high school	12 (0.26%)
High school degree or equivalent	137 (2.9%)
Some college, associate degree	883 (19.0%)
College degree	1,492 (32.1%)
Master's degree	1,438 (30.9%)
Doctorate/professional degree	687 (14.8%)
Caregiver Marital Status	
Married	3,845 (82.7%)
Divorced	391 (8.4%)
Widowed	32 (0.69%)
Separated	111 (2.4%)
Never married	127 (2.7%)
Living with partner	142 (3.1%)
Child Behavior Checklist (CBCL) T Scores	
Internalizing	48.7 (10.5)
Externalizing	45.4 (10.0)
Total problems	45.9 (10.8)

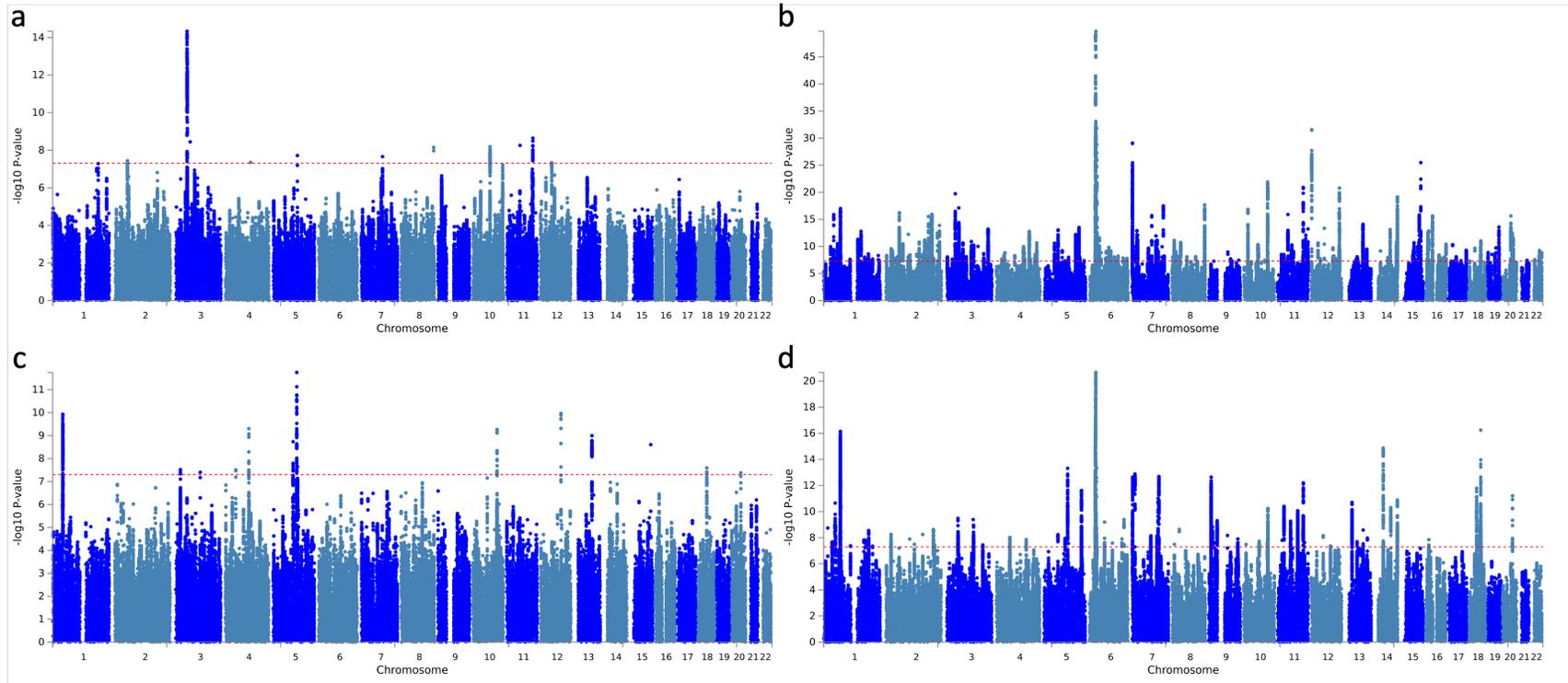
*Note.* Measures included here are for descriptive purposes. Household income, caregiver education, and marital status were not included in analyses. CBCL raw scores, rather than T scores, were used in analyses.

### Extended Data Figure 1. Genetic Correlations Among 11 Psychiatric Phenotypes



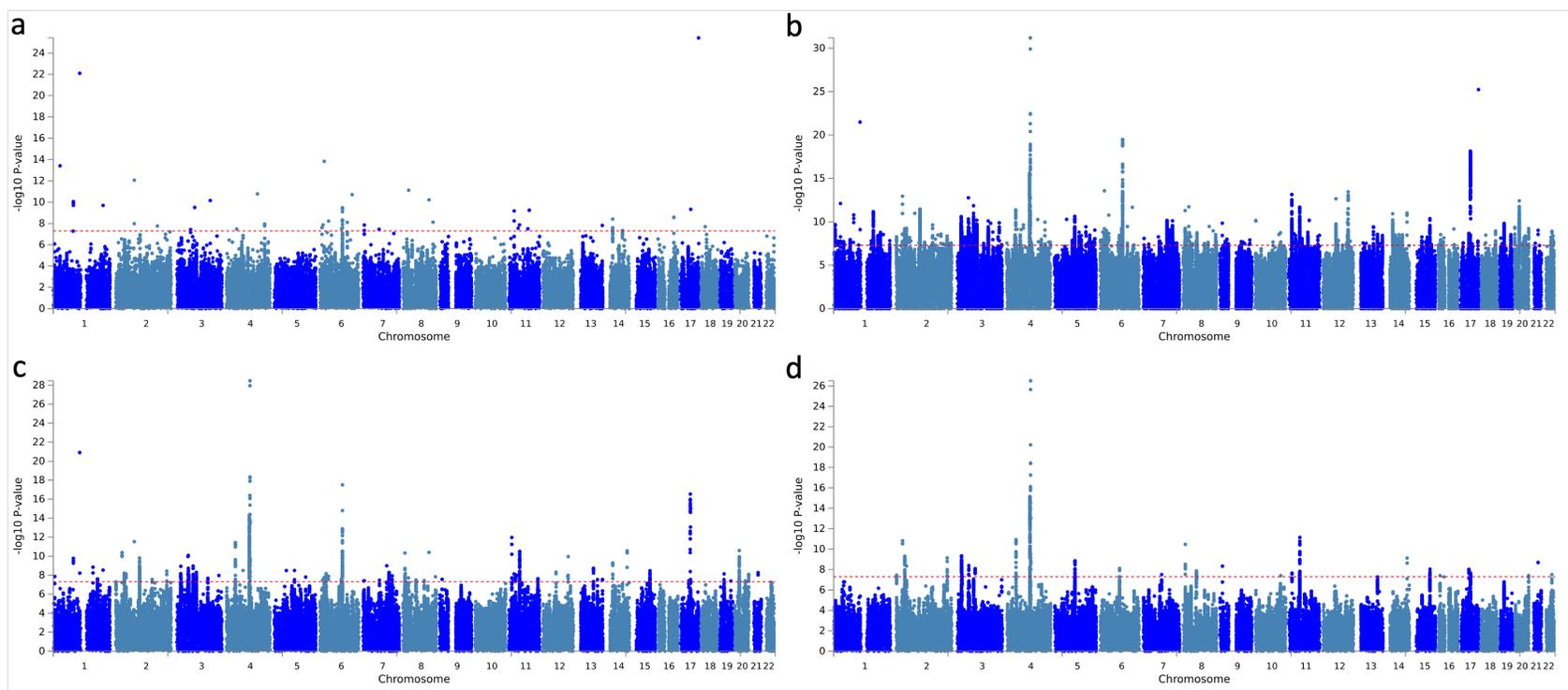
AN = Anorexia Nervosa; OCD = Obsessive-Compulsive Disorder; TS = Tourette Syndrome; SCZ = Schizophrenia; BiP = Bipolar Disorder; PAU = Problematic Alcohol Use; ADHD = Attention-Deficit/Hyperactivity Disorder; ASD = Autism Spectrum Disorder; PTSD = Post-Traumatic Stress Disorder; MDD = Major Depressive Disorder; GAD = Generalized Anxiety Disorder. Genetic correlations were estimated using the `ldsc()` function in the GenomicSEM R package.

## Extended Data Figure 2. Multivariate GWAS Manhattan Plots.



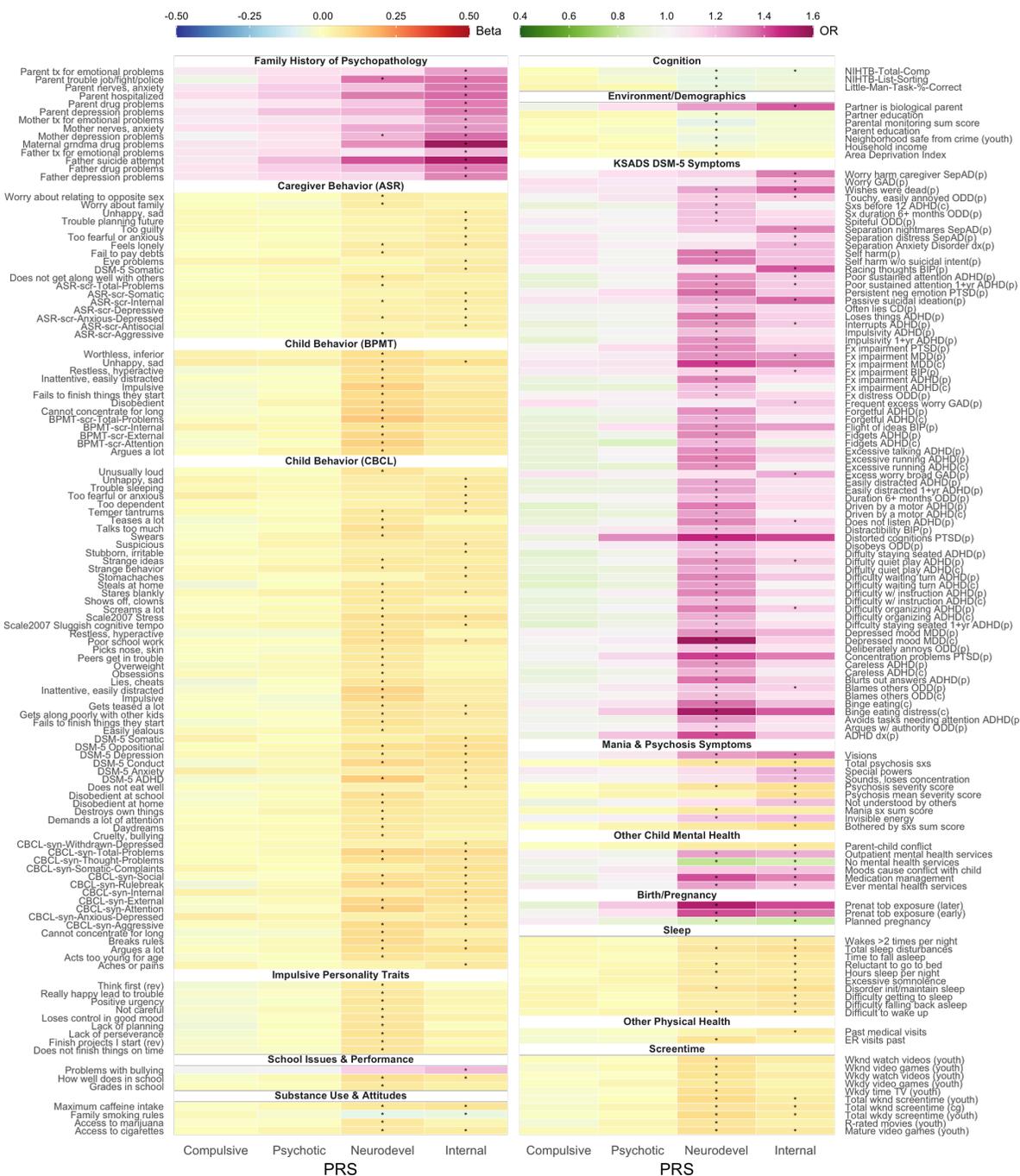
Multivariate GWAS Manhattan Plots for **a)** Compulsive, **b)** Psychotic, **c)** Neurodevelopmental, and **d)** Internalizing factors. Y-axes represent the negative  $\log_{10}$  of the SNP effect p-values. The red dotted line reflects the genome-wide significance threshold for one million tests. The x-axes show individual SNPs across chromosomes. Each dot represents a single SNP effect. Overlapping data points are not drawn. Note the different y-axes.

### Extended Data Figure 3. Multivariate GWAS Factor-Specific Q-SNPs Manhattan Plots.



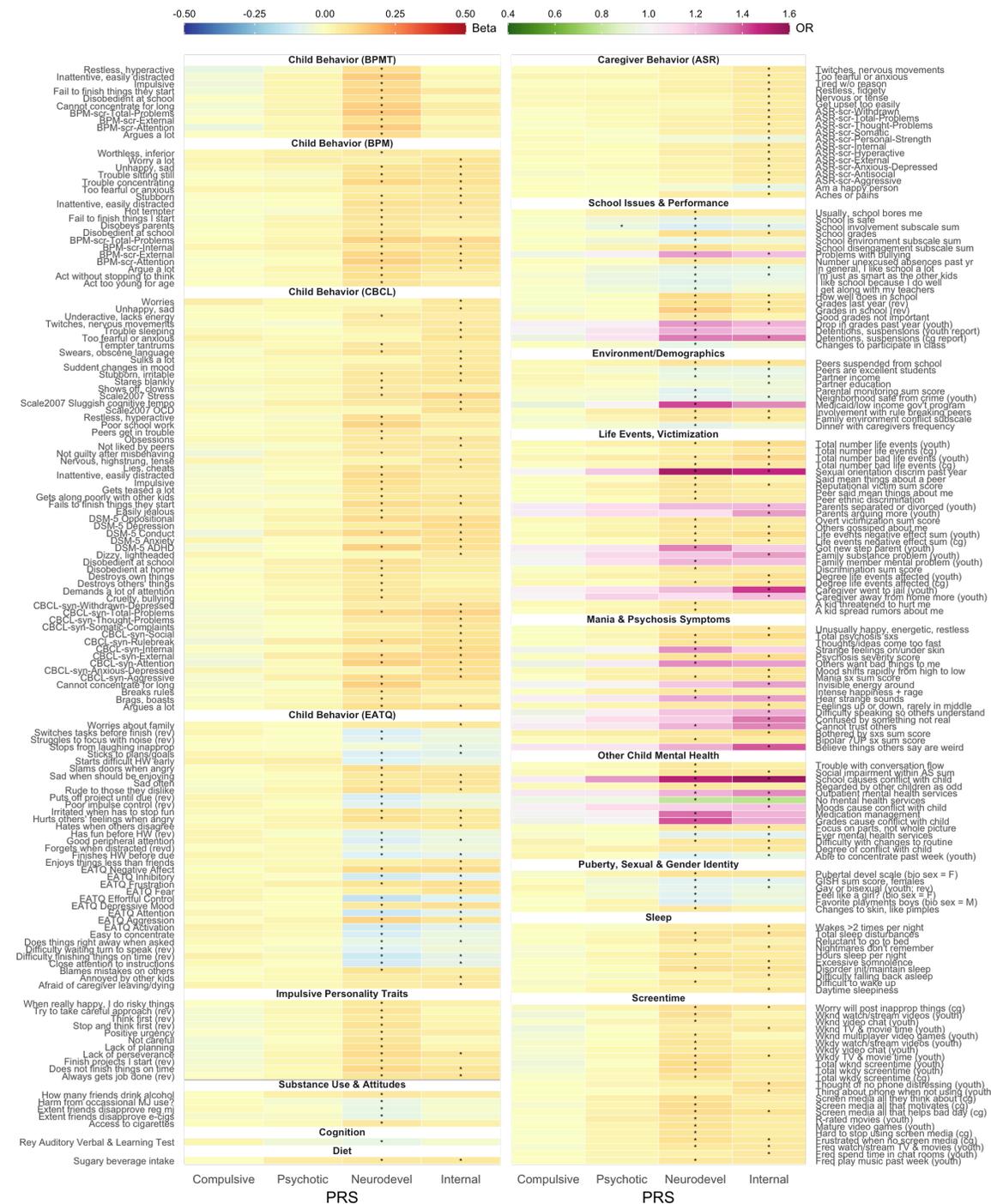
Factor-specific Q-SNP effects for **a)** Compulsive, **b)** Psychotic, **c)** Neurodevelopmental, and **d)** Internalizing factors. Y-axes represent the negative log(10) of the Q-SNP effect p-values. The red dotted line reflects the genome-wide significance threshold for one million tests. The x-axes show individual SNPs across chromosomes. Each dot represents a single Q-SNP effect. Overlapping data points are not drawn. Note the different y-axes.

### Extended Data Figure 4. Heatmap of PheWAS baseline results.



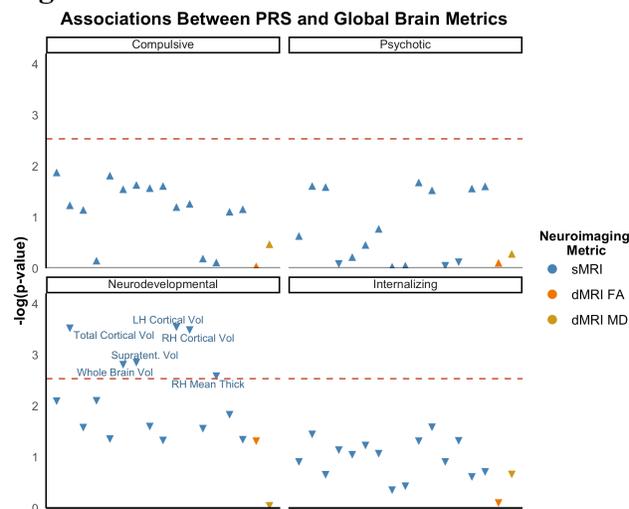
PRS are depicted along the x-axis and all phenotypes whose association with any PRS survives Bonferroni correction are labeled on the y-axes. For all phenotypes, darker colors represent larger effects. For continuous phenotypes, red colors indicate positive associations, whereas blue colors indicate negative associations. For dichotomous phenotypes analyzed with binomial generalized mixed effects models, effects sizes are represented by odds ratios (ORs), with ORs > 1 appearing pink/purple, and ORs < 1 appearing green. Asterisks show Bonferroni-corrected significant associations. Phenotypes have been grouped into domains for easier interpretation.

## Extended Data Figure 5. Heatmap of PheWAS two-year follow-up results.



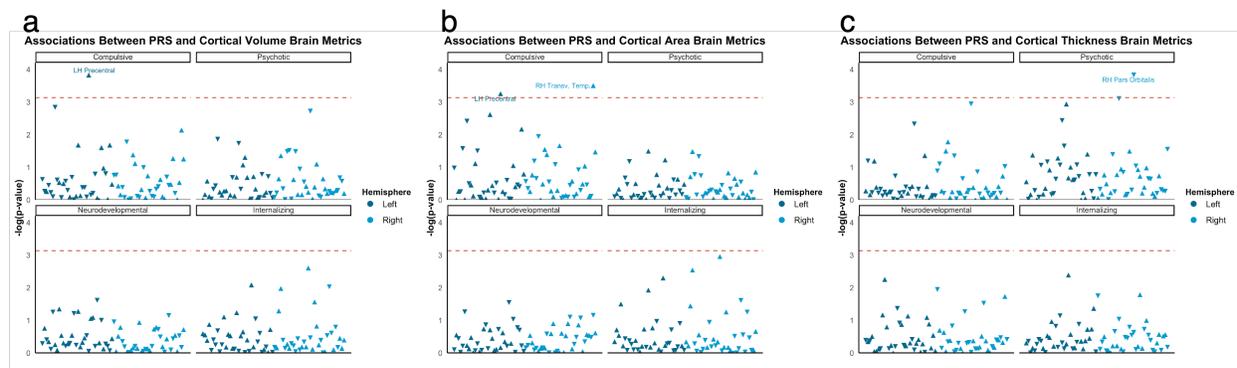
PRS are depicted along the x-axis and all phenotypes whose association with any PRS survives Bonferroni correction are labeled on the y-axes. For all phenotypes, darker colors represent larger effects. For continuous phenotypes, red colors indicate positive associations, whereas blue colors indicate negative associations. For dichotomous phenotypes analyzed with binomial generalized mixed effects models, effects sizes are represented by odds ratios (ORs), with ORs > 1 appearing pink/purple, and ORs < 1 appearing green. Asterisks show Bonferroni-corrected significant associations. Phenotypes have been grouped into domains for easier interpretation.

## Extended Data Figure 6. Associations between PRS and Global Brain Metrics.



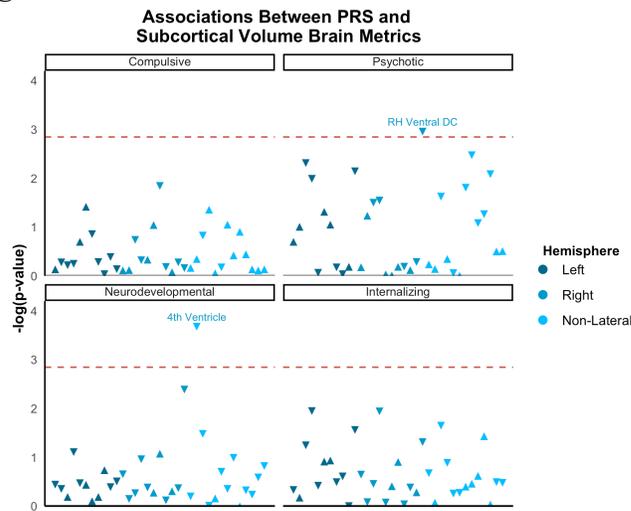
Associations between each PRS and global brain metrics. The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with the colors indicating categories and triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. LH = left hemisphere; RH = right hemisphere; supratent = supratentorial; vol = volume. Full results can be found in [Supplemental Table S13](#).

## Extended Data Figure 7. Associations between PRS and Cortical Structure.



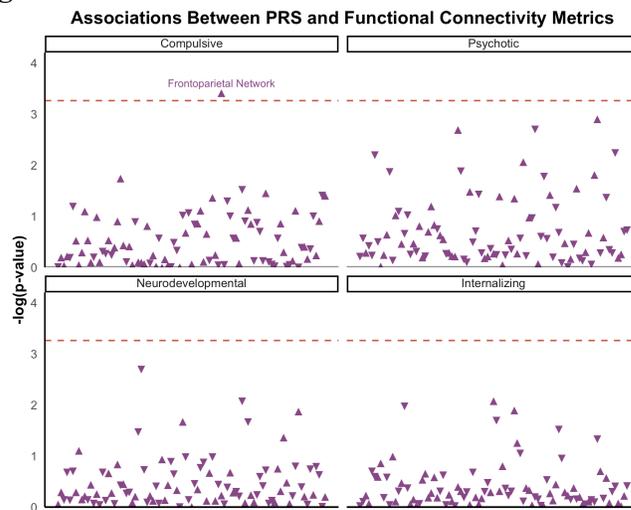
Associations between each PRS and cortical gray matter structure. **a)** Cortical Volume. **b)** Cortical Surface Area. **c)** Cortical Thickness. The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with the colors indicating hemispheres and triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. LH = left hemisphere; RH = right hemisphere; transv = transverse; temp = temporal. Full results can be found in [Supplemental Tables S14-S16](#).

### Extended Data Figure 8. Associations between PRS and Subcortical Volume.



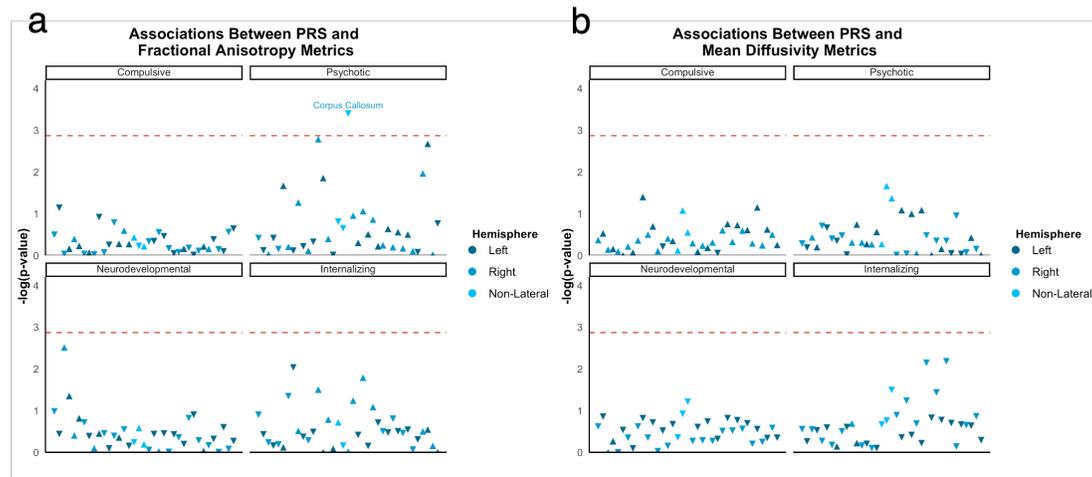
Associations between each PRS and subcortical gray matter volume. The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with the colors indicating laterality and triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. LH = left hemisphere; RH = right hemisphere; DC = diencephalon. Full results can be found in [Supplemental Table S17](#).

### Extended Data Figure 9. Associations between PRS and Functional Connectivity Metrics.



Associations between each PRS and resting state functional connectivity metrics. The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. Full results can be found in [Supplemental Table S18](#).

## Extended Data Figure 10. Associations between PRS and Diffusion Tensor Imaging Metrics.



Associations between each PRS and diffusion tensor imaging metrics. **a)** Fractional Anisotropy. **b)** Mean Diffusivity. The x-axes represent phenotypes, and the y-axis represent the  $-\log_{10}$  of the uncorrected  $p$  values of associations between each PRS and each phenotype. Each dot represents one phenotype, with the colors indicating laterality and triangle direction indicating directionality of effect. The dashed lines indicate the threshold to survive Bonferroni correction. Full results can be found in [Supplemental Tables S19-S20](#).