Supplementary Online Content

Riglin L, Collishaw S, Thapar AK, et al. Association of genetic risk variants with attention-deficit/hyperactivity disorder trajectories in the general population. *JAMA Psychiatry*. Published online November 2, 2016. doi:10.1001/jamapsychiatry.2016.2817

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. ALSPAC Genotyped Data, ALSPAC ADHD Symptoms Measure, and Additional Analyses

ALSPAC Genotyped Data

In total 9912 ALSPAC children were genotyped, of whom 8365 passed quality control. Of these, 6664 had ADHD symptoms data at two-time points and were therefore included in this study. Full genotyping details and individual exclusion criteria are described elsewhere.[1] Known autosomal variants were imputed with MACH 1.0.16 Markov Chain Haplotyping software[2, 3] using CEPH individuals from phase 2 of the HapMap project (HG18) as a reference set (release 22) resulting in a total N=2,543,887 SNPs. Dosage data were transformed from MACH output to PLINK format using fcGENE.[4] After quality control exclusions (call rate <95%, MAF <1%, HWE $P>5x10^{-7}$, $R^2 \geq .7$) there were 1,813,169 autosomal SNPs. EIGENSTRAT principal components analysis was used to generate the top 100 components after the removal of known regions of long linkage disequilibrium in the data.[5, 6] EIGENSTRAT analysis revealed no additional obvious population stratification and genome-wide analyses with other phenotypes indicate a low lambda. Previous work has included the top 10 EIGENSTRAT principal components analysis in secondary analyses[7]. We investigated whether these may need to be included in our analyses by testing for differences in these 10 principal components across our outcome variables: we found weak evidence of differences between the ADHD trajectories ($\chi^2(3)=0.05$-2.80, $p=0.42$-1.00) or between groupings of individuals using ADHD cut-points at two time points ($F(3,3644)=0.11$-2.50, $p=0.06$-$.96$). EIGENSTRAT principal components were therefore not included as covariates in our analyses.

Mental disorder risk alleles were identified from the Psychiatric Genetic Consortium (PGC) meta-analysis of case-control GWAS of ADHD (5,621 cases and 13,589 controls/pseudo-controls), schizophrenia (35,476 cases and 46,839 controls), bipolar disorder (7,481 cases and 9,250 controls) and depression (9,240 cases and 9,519 controls).[8-13] PGC SNPs were limited to those that passed an imputation quality control threshold akin to that set for the target sample (INFO score >=0.7).

Autosomal SNPs that were present in both the target and discovery sample were limited to those in relative linkage equilibrium using the --clump command in PLINK[12, 14] with an R^2 threshold of .25 and a distance threshold of 500kb, retaining SNPs with the lowest association p-value. In-line with previous work,[15] for schizophrenia PRS, only a single SNP within the extended major histocompatibility complex (MHC; chromosome 6: 25-34Mb) was included due to the high linkage disequilibrium (LD) within this region. This resulted in 153,652, 185,051, 149,512, and 157,210 clumped SNPs for ADHD, schizophrenia, bipolar disorder and depression respectively. These were used to generate polygenic risk scores using the --score command. Scores were calculated as the mean number of risk alleles weighted by effect size (log odds ratio). PRS were standardized using Z-score transformation for those with genetic data included in this study (N=6664).

ALSPAC ADHD Symptoms Measure

We used ROC curve analysis to assess the validity of the ADHD symptom subscale of the Strengths and Difficulties Questionnaire (SDQ)[16] in detecting a DSM-IV ADHD diagnosis at age 7. ADHD diagnosis was assessed using parent reports of the Development and Well-Being Assessment, a well-established research diagnostic assessment.[17] Diagnoses were generated originally using computer generated

diagnoses designed to complement clinician diagnoses and for examining risk factors.[18] Children were defined as having a diagnosis if they were in the highest computer predicted band (70% of children in this band predicted to have disorder). Using the SDQ 'abnormal' cut-point of ≥7, the area under the curve was 0.88. Of those with who had a diagnosis, around 86% reached the SDQ cut-point (sensitivity) and of those who did not have a diagnosis around 90% did not reached the SDQ cut-point (specificity); of those who reached the SDQ cut-point, 6% had a diagnosis (positive predictive value) and of those who did not reach the SDQ cut-point 99.9% did not have a diagnosis (negative predictive value), which reflect the low prevalence of ADHD diagnosis in this sample.

Additional analyses: grouping individuals using ADHD cut-points at two time-points

We grouped individuals based on parent rated Strengths and Difficulties Questionnaire[16] responses at two-ages to categorize individuals as low, childhood-limited, persistent or adolescent-onset, in-line with definitions in the main manuscript.

a) Age 4 and 17 years

Using data from age 4 and 17 years (N=4915), 83.9% of individuals were categorized as low, 10.7% childhood-limited, 1.9% persistent and 2.8% adolescent-onset (.7% excluded as they reached the cut-point at age 17 but were borderline at age 4). ADHD PRS differed across the subgroups (F(3,3677)=4.71, *p*=0.003). PRS for schizophrenia, bipolar disorder and depression did not differ by subgroup (F(3,3677)=0.43-1.05, *p*=0.37-.73). Specifically, there was evidence of higher ADHD PRS in the persistent compared to the low (B=0.34, SE=0.12, *p*=0.01). There was

evidence of higher ADHD PRS in the childhood-limited compared to the low (B=0.12, SE=0.05, *p*=0.03), as shown in Figure S2a.

b) Age 12 and 17 years

Using data from age 12 and 17 years (N=4953), 90.9% of individuals were categorized as low, 3.9% childhood-limited, 2.2% persistent and 2.4% adolescent-onset (.6% excluded as they reached the cut-point at age 17 but were borderline at age 12). ADHD PRS differed across the subgroups (F(3,3746)=7.93, *p*<0.001). PRS for schizophrenia, bipolar disorder and depression did not differ by subgroup (F(3,3746)=0.60-1.00, *p*=0.39-.62). Specifically, there was evidence of higher ADHD PRS in the persistent compared to the low (B=0.51, SE=0.11, *p*<0.001), childhood-limited (B=0.43, SE=0.15, *p*=0.004) and adolescent-onset (B=0.70, SE=0.19, *p*<0.001) subgroups, as shown in Figure S2b.

Age 7 and 17: inattentive and hyperactive-impulsive symptoms separately

Individuals were also grouped based on parent rated Strengths and Difficulties Questionnaire[16] responses at ages 7 and 17 years separately for inattentive and hyperactive-impulsive symptoms. Inattentive scores were the summed items 'Easily distracted, concentration wanders' and 'Sees tasks through to the end' (reverse coded) (possible range 0-4). Hyperactive-inattentive scores were the summed items 'Restless, overactive', 'Constantly fidgeting or squirming' and 'Thinks things out before acting' (reverse coded) (possible range 0-6). Somewhat in-line with the cut-point for abnormal scores for the total subscale being 7/10 (70% of total possible score)[16] the cut-point for inattentive and hyperactive-impulsive symptoms were selected as 3 and 4 respectively.

c) Inattentive symptoms

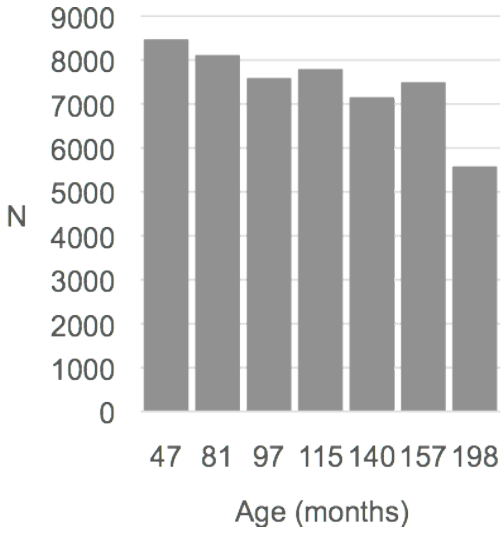
For inattentive symptoms, (N=4757), 77.2% of individuals were categorized as low, 8.3% childhood-limited, 5.6% persistent and 8.9% adolescent-onset. ADHD PRS differed across the subgroups (F(3,3623)=4.34, *p*=0.01). PRS for schizophrenia, bipolar disorder and depression did not differ by subgroup (F(3,3623)=0.18-1.92, *p*=0.12-91). Specifically, there was evidence of higher ADHD PRS in the persistent compared to the low (B=0.23, SE=0.07, *p*=0.001), as shown in Figure S2c.

d) Hyperactive-impulsive symptoms

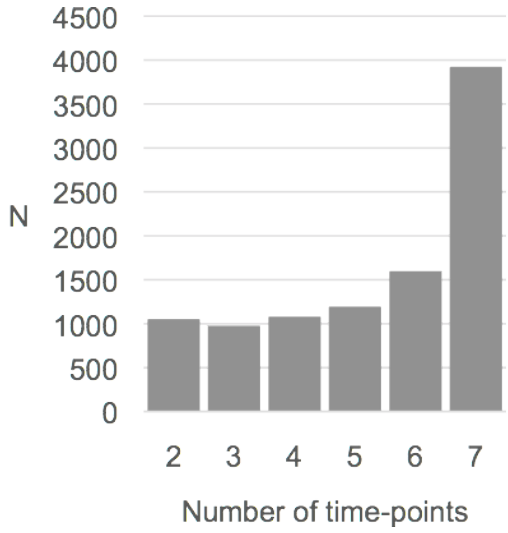
For hyperactive-impulsive symptoms (N=4535) 82.2% of individuals were categorized as low, 11.8% childhood-limited, 2.9% persistent and 3.2% adolescent-onset. ADHD PRS differed across the subgroups (F(3,3452)=3.06, *p*=0.03). PRS for schizophrenia, bipolar disorder and depression did not differ by subgroup (F(3,3452)=0.30-1.62, *p*=0.18-.83). Specifically, there was evidence of higher ADHD PRS in the persistent compared to the low (B=0.25, SE=0.10, *p*=0.02) subgroup, as shown in Figure S2d.

eFigure 1. Number of Individuals With ADHD Symptom Data at Each Time Point

a) Individuals with ADHD symptom data at each time point



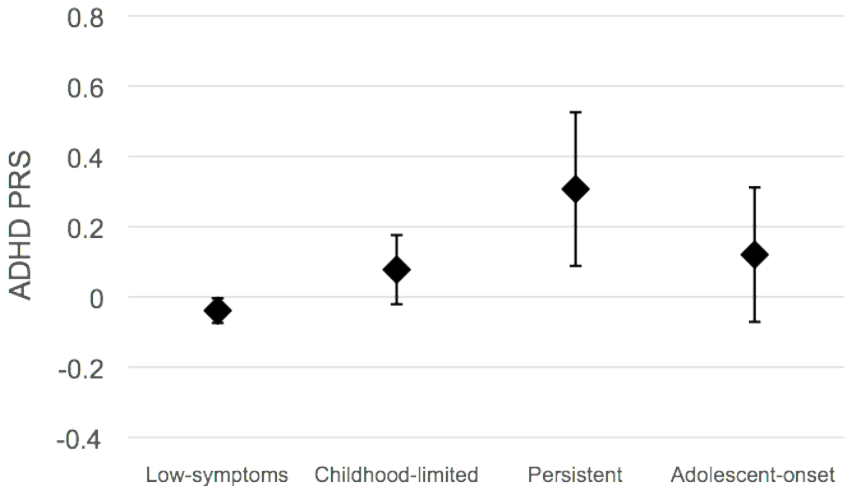
b) Number of time-points individuals had ADHD symptom data available



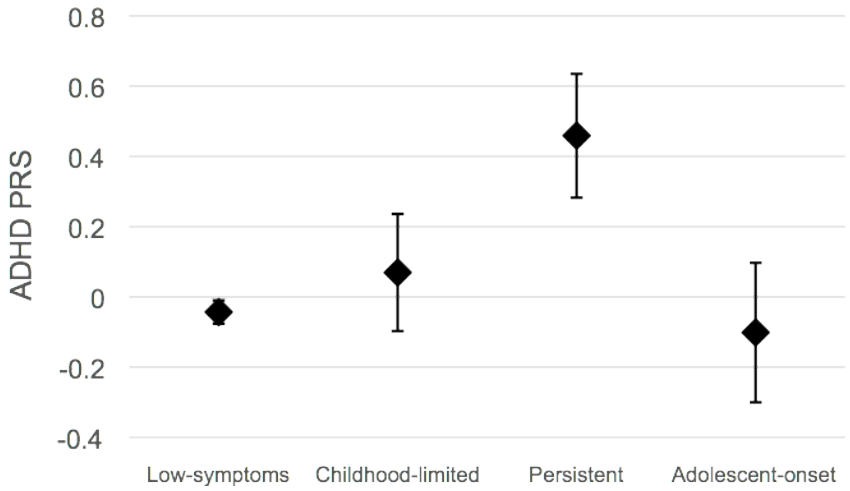
NB. Individuals with ADHD symptom data at fewer time points were less likely to be in the low class (mean number of time-points 5.38) compared to the intermediate, childhood-limited or persistent classes (mean number of time-points = 5.14, 5.19 and 4.96 respectively) ($\chi^2(1)=3.72\text{-}14.44$, $P < .001\text{-}.054$).

eFigure 2. Mean ADHD Polygenic Risk Score (PRS) by Subgroup, Based on 2 Time Points with 95% CI Error Bars

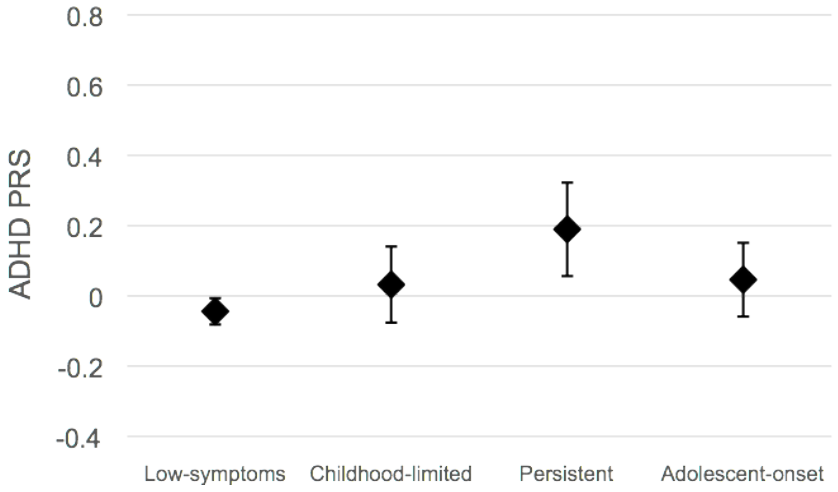
a) Age 4 and 17 years



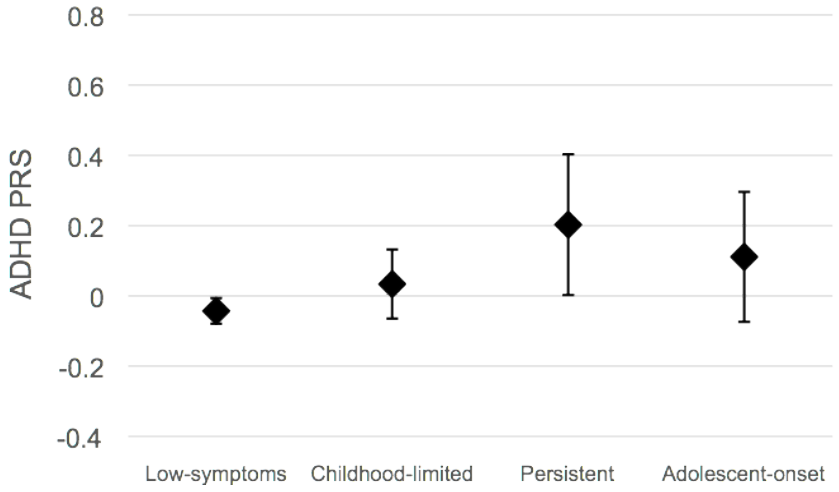
b) Age 12 and 17 years



c) Age 7 and 17 years: Inattentive symptoms



d) Age 7 and 17 years: Hyperactive-impulsive symptoms



References

1. 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. *Biological Psychiatry.* 2014;76(8):664-671.
2. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annual Review of Genomics and Human Genetics.* 2009;10:387.
3. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology.* 2010;34(8):816-834.
4. Roshyara NR, Scholz M. fcGENE: A Versatile Tool for Processing and Transforming SNP Datasets. *PLoS ONE.* 2014;9(7):e97589.
5. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics.* 2006;38(8):904-909.
6. Price AL, Weale ME, Patterson N, et al. Long-range LD can confound genome scans in admixed populations. *American Journal of Human Genetics.* 2008;83(1):132-135; author reply 135-139.
7. 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. *Journal of Child Psychology and Psychiatry.* 2015;56(6):648-656.
8. Cross-Disorder Group of the Psychiatric Genomics C. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics.* 09//print 2013;45(9):984-994.
9. Neale BM, Medland SE, Ripke S, et al. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2010;49(9):884-897.
10. Yang L, Neale BM, Liu L, et al. Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: Genome‐wide association study of both common and rare variants. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2013;162(5):419-430.
11. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-427.
12. Sklar P, Ripke S, Scott LJ, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics.* 2011;43(10):977-983.
13. Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry.* 2013;18(4):497-511.
14. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics.* 2007;81(3):559-575.
15. Jones HJ, Stergiakouli E, Tansey KE, et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry.* Published online January 27, 2016 2016.

16. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry.* Nov 2001;40(11):1337-1345.

17. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry.* 2000;41(05):645-655.

18. Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Social Psychiatry and Psychiatric Epidemiology.* 2011;46(6):521-532.