Title. Schizophrenia risk alleles are associated with neurodevelopmental outcomes in childhood

Lucy Riglin, PhD[1]

Stephan Collishaw, PhD[1]

Alexander Richards, PhD[1]

Ajay K Thapar, MRCGP[1]

Barbara Maughan, PhD[2]

Michael C O'Donovan, FRCPsych[1]

Anita Thapar, FRCPsych[1*]

[1]Division of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, UK; [2] MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Kings College London, UK.

[*]Corresponding author. Division of Psychological Medicine and Clinical Neurosciences, Cardiff University School of Medicine, Hadyn Ellis Building, Maindy Road, Cathays, Cardiff CF24 4HQ. Tel: +442920 6688325. Email: thapar@cardiff.ac.uk.

Research in context

Evidence before this study

We searched PubMed for articles published in the previous 5 years (on August 24, 2016) for the terms ((“schizophrenia” OR “psychosis” OR “psychotic”) AND (“child” or “adolescent”) AND (“antecedents” OR “genetic” OR “polygenic risk scores”) and ("review")); no language restrictions were imposed. We identified two reviews of childhood antecedents to adult mental health including schizophrenia. High-risk follow-up, retrospective and population studies have observed that although schizophrenia onset typically occurs after puberty, illness is commonly preceded by observable childhood neurodevelopmental impairments that can also be viewed as traits in the general population. Schizophrenia genetic liability, as indexed by polygenic risk scores, has been found to contribute to post-pubertal mental health problems.

Added value of this study

This study adds value to previous findings in suggesting that schizophrenia genetic liability, indexed by genetic risk scores that were generated from a sample of adults with the disorder, impacts upon childhood neurodevelopment, emotional problems and behavior in the general population as early as age 4 years.

Implications of all the available evidence

Schizophrenia polygenic risk scores are associated with elevated levels of neurodevelopmental and mental health problems in the general population from early childhood to adult life. Schizophrenia genetic risk may manifest as symptoms that do not resemble psychosis.

Abstract.

Background. Schizophrenia typically onsets after puberty but is commonly preceded by observable childhood neurodevelopmental impairments. It is unknown if these childhood antecedents index genetic liability. We used polygenic risk scores (PRS) derived from a patient discovery sample as indicators of schizophrenia genetic liability. Our aim was to identify the early childhood manifestations of this liability in a UK population-based cohort.

Method. Data were primarily analyzed using regression-based analyses in the Avon Longitudinal Study of Parents and Children (ALSPAC). PRS were generated from a published Psychiatric Genomics Consortium genome-wide association study. Outcomes were childhood (age 4-9 years) dimensional measures in four developmental domains (12 indicators were explored): cognition/learning, social/communication, emotion/mood regulation and behavior (N=5100-6952).

Outcomes. At age 7-9 years schizophrenia PRS showed associations with lower performance IQ (β=-0·056, OR=1.13), poorer social understanding (β=-0·032, OR=1·06), worse language intelligibility/fluency (β=-0·032, OR=1·10), irritability (β=0·032, OR=1·07) and headstrong behavior (β=0·031, OR=1·08). Schizophrenia PRS also predicted social and behavioral impairments as early as age 4 years.

Interpretation. Childhood cognitive, social, behavioral and emotional impairments, implicated as antecedents to schizophrenia in high-risk, developmental studies, may represent early manifestations of genetic liability.

Funding. This work was supported by the Medical Research Council (MR/M012964/1).

Keywords: ALSPAC; Child; Schizophrenia; Genetics.

Schizophrenia risk alleles are associated with neurodevelopmental outcomes in childhood

Many mental disorders have pre-pubertal origins (1). Although schizophrenia typically onsets after puberty (1, 2), high-risk, longitudinal and retrospective studies show that full-blown disorder is commonly preceded by impairments that manifest earlier in development (1). Childhood neurodevelopmental impairments involving cognition/learning, social/communication difficulties, emotion/mood dysregulation and behavior problems are known to predate the onset of schizophrenia (1-3), but it is not yet known whether these childhood antecedents index genetic liability for the disorder (1).

Schizophrenia is highly heritable; although its genetic architecture is not fully resolved, a substantial amount of the genetic variance is explained by common risk alleles (minor allele frequency ≥1%)(4). Composite polygenic risk scores (PRS), derived from these risk alleles are now considered useful indices of genetic liability (5) and provide biologically valid indicators of disease risk for research (6). Moreover, there is emerging evidence that schizophrenia PRS predict cognitive ability and post-pubertal psychopathology including negative symptoms, but not psychotic-like symptoms, in the general population (7). Thus, before the typical age of illness onset, schizophrenia genetic liability may manifest as symptoms that do not resemble psychosis. Identifying the impact of schizophrenia risk alleles on pre-pubertal, developmental characteristics in population-based samples may help to identify and better understand the early origins of this disorder and the initial manifestations of genetic liability.

This study set out to investigate the relationships between genetic risk for schizophrenia, as indexed by PRS, and pre-pubertal developmental impairments assessed at ages 7-9 years in a large population-based cohort. We focused on developmental domains that have previously been implicated in the antecedent literature for schizophrenia (1-3): (a) cognition and learning, (b) social/communication problems, (c) emotion/mood dysregulation and (d) behavior difficulties. The aim of this study was to test the hypothesis that schizophrenia genetic liability impacts on early childhood development across these domains (and that they thus represent trait liabilities) in a population-based birth cohort, the Avon Longitudinal Study of Parents and Children (ALSPAC). We also investigated whether associations extended to an earlier age (age 4). We hypothesized that schizophrenia PRS, a disorder considered by many as neurodevelopmental in origin, would impact on all of the pre-pubertal domains that in high-risk samples have been reported to be "antecedent features".

Method

ALSPAC sample

The Avon Longitudinal Study of Parents and Children is a well-established prospective, longitudinal birth cohort study. The enrolled core sample consisted of 14,541 mothers living in Avon, England, who had expected delivery dates of between 1st April 1991 and 31st December 1992. Of these pregnancies 13,988 children were alive at 1 year. When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally, resulting in an additional 713 children being enrolled. The resulting total sample size of children who were alive at 1 year was N=14,701. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Following quality control, genotype data were available for 8365 children. Phenotype data were available for between 5100-6952 individuals depending on the measures. Full details of the study, measures and sample can be found elsewhere (8, 9) (see http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary and the online data supplement).

Polygenic risk scores

Genotyping details, as well as full methods for generating the PRS, are given in the online data supplement. In brief, PRS were generated as the weighted mean number of disorder risk alleles in approximate linkage equilibrium ($R^2<0{\cdot}25$), defined in previously published GWAS, using standard procedures (6). In ALSPAC these were derived from dosage data of 1,813,169 imputed autosomal SNPs (see the online data supplement for imputation details). Risk alleles were identified as those associated with case-status in the Psychiatric Genetic Consortium (PGC) analyses (35,476 cases and 46,839 controls) (https://www.med.unc.edu/pgc/results-and-downloads) (4) at a threshold of $p<0{\cdot}05$, as this threshold maximally captures phenotypic variance for this disorder (4). Associations across a range of p-thresholds are shown in Figure S1 in the online data supplement.

Phenotypic data - outcome variables

Primary outcome variables were assessed at ages 7-9 years. Descriptive information, including correlations between variables, is included in the online data supplement Table S1.

1) *Cognition and learning* included*:* a) inattention, b) reading ability, c) verbal IQ, and d) performance IQ. Inattention was assessed using nine ADHD items from the parent-rated Development and Well-Being Assessment (DAWBA; 10), a structured diagnostic assessment widely used in child mental health surveys (individual item range 0-2). Reading ability was measured using the Wechsler Objective Reading Dimensions (11) and verbal/performance IQ using the Wechsler Intelligence Scale for Children (12); these were standardized using a Z-score transformation.

2) Social/communication included: a) social understanding, b) language intelligibility and fluency, and c) pragmatic language. Social understanding was measured by four items from the Social and Communication Disorders Checklist (SCDC; 13, 14) (possible range 0-8; reverse scored - higher scores indicate greater social understanding). Measures of

intelligibility/fluency and pragmatic language were derived from the Children's Communication Checklist (CCC; 15), and were comprised of 11 and 38 items respectively (possible ranges 16-38 and 86-162).

3) Emotion/mood regulation was assessed using the DAWBA. Scores were computed for: a) irritability, which included temper tantrums, being touchy/easily annoyed and being angry and resentful (16), and b) anxiety, composed by summing six generalized anxiety items.

4) *Behavior* included observable behaviors: a) headstrong behavior, b) aggression, and c) activity/impulsiveness, all measured by DAWBA items. Headstrong items included arguing with grown-ups, ignoring rules/refusing to do as told, doing things to annoy other people on purpose and blaming others for his/her own mistakes/bad behavior (16). Aggression included starting fights and bullying/threatening people. Activity/impulsiveness was measured by nine DAWBA ADHD items.

Age 4 years

DAWBA data were not collected prior to age 7 years but related questionnaire measures were available at age 4 years. The Strengths and Difficulties Questionnaire (SDQ; 17) is a brief, widely used questionnaire designed to assess different domains of children's mental health. It was completed by parents when children were aged 4. SDQ data were also available at age 7 years and are presented in the supplementary material to allow comparison across ages using the same measure. The subscales (each comprising 5 items, individual item range 0-2) included prosocial behavior (e.g. considerate of other people's feelings), emotional problems (e.g. many worries) and conduct (behavior) problems (e.g. often lies or cheats). The conduct problems subscale includes an irritability item (temper tantrums), which was analyzed separately as this has been found to be an indicator of emotion/mood dysregulation (e.g. see 16). All descriptive information is included in the online data supplement and Table S2, with associations with primary measures in Table S3.

Statistical analysis

Initial univariate regression analyses involved one predictor (schizophrenia PRS) and multiple dimensional outcomes (12 phenotypic measures within the 4 domains). We used a false discovery rate (18) to correct for multiple testing in our primary analyses using R (19). Given that our phenotypic measures are correlated, traditional methods of correcting for multiple testing, such as the Bonferroni method, would be overly conservative (20). Analyses were conducted in Mplus using a robust maximum likelihood parameter estimator and full information maximum likelihood estimation where data were present for at least one outcome variable (21). We also generated odds ratios for dichotomized versions of the outcome indicators (≥1 symptom for DAWBA and bottom 10% for reading, IQ, and social/communication variables, in line with previous work (e.g. 7) (percentages in each category are given in Supplementary Table 1).

We further considered the possibility of potential confounders (child gender, social class); the sample is ancestrally homogeneous (see supplementary material).

Role of the funding source

The study sponsor played no role in the study design or collection, analysis, and interpretation of data, writing of the report or decision to submit the paper for publication.

Results

Univariate associations for age 7-9 phenotypes are shown in Table 1. In the cognition/learning domain we observed an association between schizophrenia PRS and lower performance IQ, but not with inattention, reading or verbal IQ. Within the social/communication domain, we observed associations with poorer social understanding and lower language intelligibility/fluency, but not with pragmatic language. Within the emotion/mood regulation domain we observed an association between schizophrenia PRS and irritability, but not with anxiety. Within the behavior domain, we found an association specifically with headstrong behavior, but not with aggression or activity/impulsivity. These associations with schizophrenia PRS were significant after correcting for multiple testing and false discovery rate adjusted *p*-values are reported in Table 1.

When controlling for child gender and social class, associations with schizophrenia PRS remained for performance IQ, intelligibility/fluency and headstrong behavior, but not for social understanding or irritability (see discussion and Supplementary Table S5).

Effect sizes were in keeping with previous findings for PRS in epidemiological research (7, 22); adopting the approach used by Kendler (23), we estimated that individuals in the top 2·5% for schizophrenia PRS would be at roughly a 12-26% increased risk of high versus low scores for the different phenotypes.

Table 2 shows the results of secondary analyses which examined associations between schizophrenia PRS and age 4 SDQ outcomes. Associations with schizophrenia PRS were found for social difficulties and behavior problems, but not for emotion/mood regulation. At age 7 years, findings for the SDQ sub-scales were similar to those for the DAWBA data with associations observed for social difficulties, emotion/mood regulation and behavior problems, presented in the online data supplement Table S4.

Discussion

This study set out to investigate the relationship between schizophrenia risk alleles, as indexed by PRS derived from a sample of patients with disorder, and pre-pubertal developmental impairments in a population-based sample of children. The clinical manifestations of schizophrenia are typically post-pubertal, but the disorder is considered by many to have a strong neurodevelopmental component and to be preceded by developmental deficits. It has not been known whether these early childhood difficulties (retrospectively recalled by those with schizophrenia and observed in high-risk studies [e.g. offspring of parents with schizophrenia]), are early manifestations of genetic liability. In this population-based sample, schizophrenia PRS showed associations with pre-pubertal performance IQ, social/communication difficulties, emotion/mood dysregulation and behavior problems. Cognitive, language and social impairments as well as emotional and behavioral difficulties have been well documented in children who went on to develop schizophrenia in high-risk follow-up, case-control and follow-back studies (1, 3). However, the majority of children who show such deficits do not later develop disorder, and whether these are indicators of genetic liability has been questioned (1). Our work suggests that pre-pubertal lower performance IQ, poorer social understanding and language intelligibility/fluency, irritability and headstrong behavior, may be early manifestations of schizophrenia genetic liability. This now requires testing. Our results also highlight that schizophrenia PRS contribute to traits that are observable in the general population from a very early age, many years before the onset of any adult forms of psychopathology *per se*. Given that the prevalence of schizophrenia in the general population is low, the findings suggest that these pre-pubertal features represent indices of liability rather than an illness prodrome.

Regarding the specific developmental domains, we found evidence of association with schizophrenia genetic risk scores for performance IQ. Other measures of cognitive ability were not available, although genetic overlap between schizophrenia and pre-pubertal performance IQ specifically has been identified by previous work (24). Our work extends these findings by suggesting that links are not generalized to other aspects of cognition/learning including inattention, reading and verbal IQ (that are predicted by ADHD PRS despite ADHD genetic discovery samples being much smaller and thus less well powered than those for schizophrenia (22)).

Schizophrenia risk was also associated with social/communication difficulties as early as age 4. Social impairments and communication skills have received less attention than cognitive features as possible early antecedents of mental disorders. Interestingly some of these social/communication difficulties could be regarded as similar to negative symptoms of schizophrenia that show post-pubertal associations with schizophrenia PRS (7). Our findings suggest that these domains of development that impact on early socialization, such as prosocial behavior, may also be manifestations of genetic liability to schizophrenia.

While childhood emotional problems are most commonly considered precursors to mood disorders, a recent review suggests schizophrenia spectrum disorders are preceded by emotional problems in middle childhood (25) and recent evidence has found an association between schizophrenia genetic risk scores and post-pubertal anxiety disorder at age 16 in the general population (7). Our findings were mixed for emotional problems; associations were observed for irritability but not for our diagnostic measure of generalized anxiety symptoms.

Associations between schizophrenia PRS and behavioral problems were also found as early as age 4. While behavioral problems are often considered largely environmentally driven, our finding is consistent with a neurodevelopmental component to early-onset behavioral problems (26), which for some may index genetic liability to adult onset mental disorders.

Our findings should be considered in light of some limitations. First, while our four developmental domains were conceptually selected *a priori* based on the antecedent literature (1-3), our analysis involved multiple testing, and although we attempted to adjust for this, we cannot rule out false positive findings and replication is advisable for any genetic finding. In addition, our outcome variables are inter-correlated. Further investigation into the genetic correlation between the different childhood traits we have identified as associated with schizophrenia PRS will be important to test in the future (e.g. 27).

Another limitation is that DAWBA data, cognitive and language data were not available before the age of 7 years. Correlations between primary (interview) and secondary (questionnaire) measures in ALSPAC were modest, and the extent to which these reflect the same underlying construct is unclear. Thus, the questionnaire findings do not represent an internal replication. Further, after correction for multiple testing, although the associations between social understanding, intelligibility/fluency, irritability and headstrong behavior and schizophrenia PRS were significant, they represent only weak evidence of an association with the outcome.

However, we provide novel evidence that certain pre-pubertal features found to be antecedents to schizophrenia index genetic liability and highlight the importance of further investigating these features at a very early age in high-risk groups. Future investigation into the robustness of specific findings will be important and our population findings provide specific hypotheses to now test in high-risk longitudinal samples.

Another limitation is that our target sample is a longitudinal birth cohort study that inevitably faces issues of non-random attrition. This likely resulted in a retained sample with lower PRS and fewer developmental impairments - which may have resulted in underestimated associations between PRS and pre-pubertal characteristics. Finally, although PRS provide a useful indicator of genetic liability (see 23), they are not a recommended method for explaining a substantial amount of phenotype variance of population traits that may only be weakly correlated with risk of disorder – indeed, schizophrenia PRS only explain a small amount of the variance in childhood neurodevelopment in our analyses and some associations did not remain when controlling for social class, although schizophrenia PRS were not associated with social class. The effect sizes of our associations are small but typical for this kind of work using PRS (e.g. 7, 22).

The present study indicates that for schizophrenia liability such characteristics may be present as early as age 4 years old. An important future research goal would be to distinguish between early manifestations of liability that reflect pleiotropy and those that represent developmental impairments that are causally associated with schizophrenia. Intervention studies are likely to be useful here. Pleiotropic effects may provide insight into cross-diagnostic nosology and trans-diagnostic processes (28), while causal factors may help to inform interventions that promote resilience to future impairment (29).

Disclosures and acknowledgments.

All authors report no competing interests. We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (102215/2/13/2) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. This work was supported by the Medical Research Council (MR/M012964/1).

Authors' contributions

AT, MO'D, SC and BM contributed to the initial study design. All authors contributed to the manuscript writing, literature search and final approval of the manuscript. LR, AKT and AR contributed to data analyses. All authors contributed to data interpretation.

References

1. Rutter M, Kim-Cohen J, Maughan B. Continuities and discontinuities in psychopathology between childhood and adult life. Journal of Child Psychology and Psychiatry. 2006;47(3-4):276-95.
2. Pine DS, Fox NA. Childhood antecedents and risk for adult mental disorders. Annu Rev Psychol. 2015;66:459-85.
3. Fryers T, Brugha T. Childhood determinants of adult psychiatric disorder. Clinical Practice & Epidemiology in Mental Health. 2013;9(1).
4. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421-7.
5. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet. 2013;381(9875):1371-9.
6. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009;460(7256):748-52.
7. Jones HJ, Stergiakouli E, Tansey KE, et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. JAMA Psychiatry. 2016.
8. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: The 'Children of the 90s'-the index offspring of the Avon Longitudinal Study of Parents and Children. International journal of epidemiology. 2013;42(1):111-27.
9. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. International journal of epidemiology. 2013;42(1):97-110.
10. 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):645-55.
11. Rust J, Golombok S, Trickey G. WORD, Wechsler objective reading dimensions manual: Psychological Corporation; 1993.
12. Wechsler D, Golombok S, Rust J. WISC-III UK Wechsler Intelligence Scale for Children: UK Manual. Sidcup, UK: The Psychological Corporation. 1992.
13. Mandy W, Skuse D, Steer C, St Pourcain B, Oliver BR. Oppositionality and socioemotional competence: interacting risk factors in the development of childhood conduct disorder symptoms. Journal of the American Academy of Child & Adolescent Psychiatry. 2013;52(7):718-27.
14. Skuse DH, Mandy WP, Scourfield J. Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. The British Journal of Psychiatry. 2005;187(6):568-72.
15. Bishop DV. Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. Journal of Child Psychology and Psychiatry. 1998;39(06):879-91.
16. Stringaris A, Goodman R. Longitudinal outcome of youth oppositionality: irritable, headstrong, and hurtful behaviors have distinctive predictions. Journal of the American Academy of Child & Adolescent Psychiatry. 2009;48(4):404-12.
17. Goodman R. The Strengths and Difficulties Questionnaire: A research note. Journal of Child Psychology and Psychiatry. 1997;38(5):581-6.

18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society Series B (Methodological). 1995,57:289-300.
19. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. URL http://www.R-project.org/.
20. Perneger TV. What's wrong with Bonferroni adjustments. BMJ: British Medical Journal. 1998;316(7139):1236.
21. Muthén LK, Muthén BO. Mplus User's Guide. Seventh ed. Los Angeles, CA: Muthén & Muthén; 1998-2012.
22. 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):648-56.
23. Kendler KS. The schizophrenia polygenic risk score: To what does it predispose in adolescence? JAMA Psychiatry. 2016;73(3):193-4.
24. Hubbard L, Tansey KE, Rai D, Jones P, Ripke S, Chambert KD, et al. Evidence of Common Genetic Overlap Between Schizophrenia and Cognition. Schizophrenia Bulletin. 2015:sbv168.
25. Laurens KR, Luo L, Matheson SL, Carr VJ, Raudino A, Harris F, et al. Common or distinct pathways to psychosis? A systematic review of evidence from prospective studies for developmental risk factors and antecedents of the schizophrenia spectrum disorders and affective psychoses. BMC Psychiatry. 2015;15(1):205.
26. Thapar A, Rutter M. Neurodevelopmental disorders. In: Thapar A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor E, editors. Rutter's Child and Adolescent Psychiatry. Sixth ed. Oxford: Wiley Press; 2015.
27. Anttila V, Bulik-Sullivan B, Finucane HK, Bras J, Duncan L, Escott-Price V, et al. Analysis of shared heritability in common disorders of the brain. bioRxiv. 2016:048991.
28. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Medicine. 2013;11(1):126.
29. Collishaw S, Hammerton G, Mahedy L, Sellers R, Owen MJ, Craddock N, et al. Mental health resilience in at-risk adolescents. Lancet Psychiatry. 2016;3(1):49-57.

Table 1. Associations between schizophrenia polygenic risk scores (PRS) and phenotypic variables at age 7-9

	Continuous outcomes				Dichotomized outcomes#
	β	SE	FDR adjusted *p*	R^2	OR(95% CI)
Cognition/learning					
Inattention	0·012	0·013	0·46	<0·001	1·01 (0·96-1·07)
Reading	-0·017	0·013	0·26	<0·001	1·04 (0·96-1·14)
Verbal IQ	-0·020	0·013	0·24	<0·001	1·05 (0·96-1·15)
Performance IQ	-0·056	0·013	0·00034	0·003	1·13 (1·04-1·23)
Social/communication					
Social understanding	-0·032	0·013	0·043	0·001	1·08 (1·00-1·17)
Intelligibility/fluency	-0·032	0·013	0·043	0·001	1·10 (1·02-1·20)
Pragmatic language	-0·003	0·013	0·82	<0·001	0·97 (0·89-1·06)
Emotion/mood regulation					
Irritability	0·032	0·013	0·043	0·001	1·07 (1·01-1·14)
Anxiety	0·022	0·014	0·22	0·001	1·02 (0·97-1·08)
Behavior					
Headstrong	0·031	0·013	0·043	0·001	1·08 (1·02-1·15)
Aggression	0·016	0·013	0·28	<0·001	1·06 (0·98-1·16)
Activity/impulsivity	0·004	0·013	0·82	<0·001	1·01 (0·96-1·06)

Highlighted where there is evidence of an association. FDR = false discovery rate # ≥1 symptom for DAWBA, bottom 10% of distribution for reading, IQ, and social/communication variables.

Table 2. Associations between schizophrenia polygenic risk scores and SDQ subscales at age 4 years

	Continuous outcomes				Dichotomized outcomes[#]	
	β	SE	*p*	R^2	OR	(95% CI)
Social/communication						
Prosocial behavior	-0·031	0·013	0·012	0·001	1·05	(0·97-1·15)
Emotion/mood regulation						
Emotional problems	0·021	0·013	0·12	<0·001	1·01	(0·93-1·10)
Irritability	0·021	0·013	0·11	<0·001	1·04	(0·96-1·11)
Behavior						
Conduct problems	0·026	0·013	0·046	0·001	1·08	(0·99-1·18)

[#] Bottom 10% of distribution for prosocial behavior; top 10% for emotional problems, irritability and conduct problems.

Supplementary material

ALSPAC data

In total 9912 ALSPAC children were genotyped, of whom 8365 passed quality control. 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 $P>5x10^{-7}$, $R^2 \geq 0{\cdot}7$) there were 1,813,169 autosomal SNPs. The ALSPAC team used EIGENSTRAT principal components analysis 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 indicated a low lambda. In-line with previous work (14) we did not include principal components in our main analyses, but ran a sensitivity analyses including the top 10 EIGENSTRAT principal components, as has been done previously (7). These are presented in Supplementary Table 6. Individuals with genetic data who were alive at one year were included in this study (N=8125).

Both genetic and phenotypic data were available for N=55253-6157 children depending on the phenotypic variable and age.

The Strengths and Difficulties Questionnaire (SDQ)

The SDQ subscales were generated using the items presented below (8). For conduct problems, we excluded the irritability item ‘often has temper tantrums or hot tempers’, which is traditionally included in this subscale, as we consider this a dimension of emotion/mood regulation. *Prosocial behavior*: considerate of other people's feelings; shares readily with other children; helpful if someone is hurt; kind to younger children; often volunteers to help others. *Emotional problems*: often complains of headaches; many worries; often unhappy, downhearted; nervous or clingy in new situations; many fears, easily scared. *Conduct problems*: generally obedient (reverse coded); often fights with other children; often lies and cheats; steals from home, school or elsewhere. *Irritability item:* often has temper tantrums or hot tempers.

Generating polygenic risk scores

Schizophrenia risk alleles were identified from the Psychiatric Genetic Consortium (PGC) meta-analysis of case-control GWAS of schizophrenia (35,476 cases and 46,839 controls) (12). 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 (13). SNPs were strand-flipped where appropriate so that the ALSPAC and PGC SNP were strand aligned. In-line with previous work (14), SNPs were clumped with an R^2 threshold of 0·25 and a distance threshold of 500kb, retaining SNPs with the lowest association p-value - 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. While clumped results are available from the PGC, in these data the retained SNPs in each ‘clump’ are not necessarily those present in ALSPAC. We therefore ran our own clumping to maximize the number of SNPs included in the polygenic risk scores, which resulted in 185,051 clumped SNPs. 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). Polygenic risk scores were standardized using Z-score transformation.

Table S1. Means, standard deviations, and correlations between primary phenotype data

Primary variables		Descriptives				Correlations											
		Mean	(SD)	%#	(N)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	Inattention	2·47	(3·70)	51·4	(2905)	1											
2.	Reading	0·02	(0·99)	9·8	(589)	-0·25	1										
3.	Verbal IQ	0·05	(1·00)	10·5	(580)	-0·19	0·57	1									
4.	Performance IQ	0·04	(0·99)	9·5	(526)	-0·19	0·35	0·49	1								
5.	Social understanding	7·15	(1·43)	12.3	(2123)	-0·43	0·13	0·10	0·09	1							
6.	Intelligibility/fluency	35·30	(1·89)	11·1	(640)	-0·22	0·26	0·20	0·12	0·16	1						
7.	Pragmatic language	150·87	(7·74)	10·4	(586)	-0·48	0·26	0·25	0·21	0·42	0·35	1					
8.	Irritability	0·49	(1·08)	24·2	(1358)	0·41	-0·07	-0·05	-0·05	-0·39	-0·10	-0·30	1				
9.	Anxiety	0·96	(1·70)	34·5	(1950)	0·23	0·02	0·02	-0·02	-0·14	-0·03	-0·18	0·30	1			
10.	Headstrong	0·74	(1·51)	29·1	(1644)	0·48	-0·11	-0·07	-0·08	-0·45	-0·08	-0·34	0·76	0·23	1		
11.	Aggression	0·13	(0·45)	9·8	(545)	0·24	-0·11	-0·09	-0·06	-0·27	-0·07	-0·21	0·34	0·12	0·36	1	
12.	Activity/impulsivity	2·41	(3·58)	52·9	(2992)	0·71	-0·21	-0·14	-0·14	-0·46	-0·15	-0·50	0·47	0·21	0·59	0·31	1

Between-variable correlations N=4539-5631, for individual variables N=5525-6037. #Used to generate odds ratios for dichotomized outcomes (≥1 symptom for DAWBA, bottom 10% of distribution for reading, IQ and social/communication variables, in-line with previous work (e.g. 14)).

Table S2. Means, standard deviations, and correlations between secondary phenotype data (SDQ)

	Age 4				Age 7			
	1.	2.	3.	4.	1.	2.	3.	4.
1. Prosocial behavior	1				1			
2. Emotional problems	-0·11	1			-0·13	1		
3. Irritability	-0·21	0·22	1		-0·23	0·25	1	
4. Conduct problems	-0·36	0·20	0·35	1	-0·41	0·22	0·39	1
Mean	7·06	1·45	0·82	1·11	8·19	1·50	0·59	0·99
(SD)	(1.97)	(1.52)	(0.67)	(1.01)	(1.75)	(1.66)	(0.68)	(1.03)

Between-variable correlations N=6134-6143 and 5702-5707 for age 4 and age 7 respectively. For individual variables N=6143-6157 and 5716-5737 respectively.

Table S3. Correlations between primary dimensions and secondary questionnaire (SDQ) measure (age 7)

Primary dimension (domain)	Prosocial behavior	Emotional problems	Irritability	Conduct problems
Social understanding (social/communication)	0.32	-0·16	-0·23	-0·32
Irritability (emotion/mood regulation)	-0·25	0·23	0·37	0·34
Headstrong (behavior)	-0·28	0·18	0·33	0·43

N=5048-5109. Primary dimensions included where primary analyses found evidence of association with schizophrenia polygenic risk scores and where a secondary measure of the same domain was available (i.e. performance IQ and intelligibility/fluency were not included as there was no secondary measures of cognition and communication).

Table S4. Associations between schizophrenia polygenic risk scores and SDQ subscales at age 7 years

	Continuous outcomes				Dichotomized outcomes#	
	β	SE	*p*	R^2	OR	(95% CI)
Social/communication						
Prosocial behavior	-0·036	0·014	0·0083	0·001	1·14	(1·04-1·24)
Emotion/mood regulation						
Emotional problems	0·027	0·013	0·043	0·001	1·10	(1·01-1·18)
Irritability	0·030	0·013	0·024	0·001	1·12	(1·03-1·21)
Behavior						
Conduct problems	0·035	0·013	0·0084	0·001	1·05	(0·95-1·16)

Bottom 10% of distribution for prosocial behavior; top 10% for emotional problems, irritability and conduct problems.

Table S5. Associations between schizophrenia polygenic risk scores and phenotypic variables at age 7-9, controlling for child gender and parental income

	Continuous outcomes			Dichotomized outcomes#	
	β	SE	*p*	OR	(95% CI)
Cognition/learning					
Inattention	-0·001	0·015	0·96	1·00	(0·94-1·07)
Reading	-0·020	0·015	0·19	1·11	(1·00-1·24)
Verbal IQ	-0·007	0·015	0·63	1·04	(0·93-1·15)
Performance IQ	-0·046	0·016	0·0032	1·11	(1·00-1·24)
Social/communication					
Social understanding	-0·026	0·016	0·10	1·05	(0·98-1·12)
Intelligibility/fluency	-0·032	0·015	0·031	1·11	(1·01-1·23)
Pragmatic language	0·004	0·015	0·81	0·96	(0·86-1·06)
Emotion/mood regulation					
Irritability	0·021	0·015	0·18	1·06	(0·99-1·14)
Anxiety	0·017	0·017	0·33	1·00	(0·94-1·07)
Behavior					
Headstrong	0·034	0·015	0·023	1·10	(1·03-1·18)
Aggression	0·003	0·015	0·84	1·03	(0·93-1·14)
Activity/impulsivity	0·000	0·015	0·99	1·00	(0·94-1·07)

Gender coded 0=female, 1=male, income assessed as the average household income band, including social benefits, each week when the child was on a ten-point scale when the child was 134 months old. Schizophrenia

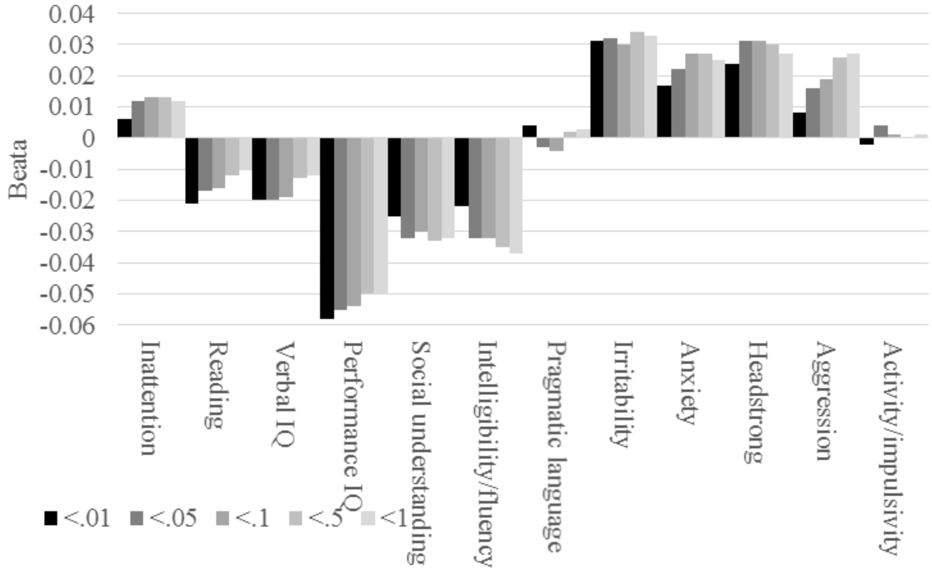
PRS was not associated with sex (OR=1.02, 95% CI = 0.98-1.06, p=0.41) or income (OR=1.02, 95% CI = 0.97-1.08, p=0.38) [#] ≥1 symptom for DAWBA, bottom 10% of distribution for reading, IQ, and social/communication variables.

Table S6. Associations between schizophrenia polygenic risk scores and phenotypic variables at age 7-9, controlling for 10 EIGENSTRAT population stratification covariates

	Continuous outcomes			Dichotomized outcomes[#]	
	β	SE	*p*	OR	(95% CI)
Cognition/learning					
Inattention	0·013	0·013	0·34	1·01	(0·96-1·07)
Reading	-0·019	0·013	0·15	1·05	(0·96-1·14)
Verbal IQ	-0·021	0·013	0·12	1·05	(0·96-1·15)
Performance IQ	-0·056	0·013	<0·0001	1·13	(1·04-1·23)
Social/communication					
Social understanding	-0·032	0·013	0·016	1·09	(1·00-1·18)
Intelligibility/fluency	-0·032	0·013	0·014	1·11	(1·02-1·20)
Pragmatic language	-0·004	0·013	0·79	0·97	(0·89-1·06)
Emotion/mood regulation					
Irritability	0·032	0·013	0·014	1·08	(1·01-1·14)
Anxiety	0·023	0·014	0·10	1·02	(0·97-1·08)
Behavior					
Headstrong	0·032	0·013	0·011	1·09	(1·03-1·15)
Aggression	0·016	0·013	0·20	1·06	(0·98-1·16)
Activity/impulsivity	0·005	0·013	0·71	1·01	(0·96-1·06)

[#] ≥1 symptom for DAWBA, bottom 10% of distribution for reading, IQ, and social/communication variables.

Figure S1. Associations between schizophrenia polygenic risk scores and phenotypic variables at age 7-9, using a range of p-value thresholds from the discovery sample



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. Biol Psychiatry. 2014;76:664-671.

2. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. Annu Rev Genomics Hum Genet. 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. Genet Epidemiol. 2010;34:816-834.

4. Roshyara NR, Scholz M. fcGENE: A Versatile Tool for Processing and Transforming SNP Datasets. PLoS One. 2014;9: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. Nat Genet. 2006;38:904-909.

6. Price AL, Weale ME, Patterson N, Myers SR, Need AC, Shianna KV, Ge D, Rotter JI, Torres E, Taylor KD. Long-range LD can confound genome scans in admixed populations. Am J Hum Genet. 2008;83:132-135; 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. J Child Psychol Psychiatry. 2015;56:648-656.

8. Goodman R. The Strengths and Difficulties Questionnaire: A research note. J Child Psychol Psychiatry. 1997;38:581-586.

9. Rutter M, Tizard J, Whitmore K: Education, health and behaviour, Longman Publishing Group; 1970.

10. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009;41:703-707.

11. Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, Phillips A, Wesley E, Parnell K, Zhang H, Drummond H. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet. 2009;41:1330-1334.

12. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421-427.

13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-575.

14. Jones HJ, Stergiakouli E, Tansey KE, et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. JAMA Psychiatry. 2016;73:221-228