



Hannigan, L. J., Askeland, R. B., Ask, H., Tesli, M., Corfield, E., Ayorech, Z. N. L., Helgeland, Ø., Magnus, P., Njølstad, P. R., Øyen, A. S., Stoltenberg, C., Andreassen, O. A., Davey Smith, G., Reichborn-Kjennerud, T., & Havdahl, K. A. (2021). Genetic liability for schizophrenia and childhood psychopathology in the general population. *Schizophrenia Bulletin*, 47(4), 1179-1189. Article sbaa193. <https://doi.org/10.1093/schbul/sbaa193>

Peer reviewed version

Link to published version (if available):
[10.1093/schbul/sbaa193](https://doi.org/10.1093/schbul/sbaa193)

[Link to publication record on the Bristol Research Portal](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Oxford University Press at <https://doi.org/10.1093/schbul/sbaa193>

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

Manuscript title: Genetic liability for schizophrenia and childhood psychopathology in the general population

Running title:

Genetic liability for schizophrenia and childhood psychopathology

Author list:

Laurie J. Hannigan^{1,2}, Ragna Bugge Askeland³, Helga Ask³, Martin Tesli^{3,4}, Elizabeth Corfield³, Ziada Ayorech^{1,2}, Øyvind Helgeland^{5,6}, Per Magnus^{7,8}, Pål Rasmus Njølstad^{5,9}, Anne-Siri Øyen^{1,11}, Camilla Stoltenberg^{10,11}, Ole A. Andreassen^{4,12}, George Davey Smith², Ted Reichborn-Kjennerud^{3,13}, Alexandra Havdahl^{1,2,3,14}

Affiliations:

¹ Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway

² MRC Integrative Epidemiology Unit, Bristol Medical School, Bristol, UK

³ Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway

⁴ NORMENT Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

⁵ KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway

⁶ Department of Genetics and Bioinformatics, Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway

⁷ Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway

⁸ Institute of Health and Society, University of Oslo, Norway

⁹ Department of Pediatrics and Adolescents, Haukeland University Hospital, Bergen, Norway

¹⁰ Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway.

¹¹ Norwegian Institute of Public Health, Oslo, Norway

¹² Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

¹³ Institute of Clinical Medicine, University of Oslo, Oslo, Norway

¹⁴ Promenta Research Center, Department of Psychology, University of Oslo, Oslo, Norway

Correspondence to:

Laurie Hannigan
Nic Waals Institute
Lovisenberg Diaconal Hospital
P.O.Box 4970 Nydalen
0440 Oslo
Norway

laurie.hannigan@bristol.ac.uk

Abstract word count	Main text word count	Number of tables	Number of figures	Supplementary Information
248	3998	2	4	Yes

Abstract (249/250 words)

Genetic liability for schizophrenia is associated with psychopathology in early life. It is not clear if these associations are time-dependent during childhood, nor if they are specific across different forms of psychopathology. Using genotype and questionnaire data on children (N = 15,105) from the Norwegian Mother, Father and Child Cohort Study (MoBa), we used schizophrenia polygenic risk scores to test developmental stability in associations with measures of emotional and behavioural problems between 18 months and 5 years, and domain specificity in associations with symptoms of depression, anxiety, conduct problems, oppositionality, inattention, and hyperactivity at 8 years. We then sought to identify symptom profiles – across development and domains – associated with schizophrenia polygenic liability. We found evidence for developmental stability in associations between schizophrenia polygenic risk scores and emotional and behavioural problems, with the latter being mediated specifically via the rate of change in symptoms ($\beta_{\text{slope}} = 0.032$; 95% CI 0.007 – 0.057). At age 8, associations were better explained by a model of symptom-specific polygenic effects, rather than effects mediated via a general psychopathology factor or by domain-specific factors. Overall, individuals with higher schizophrenia polygenic risk scores were more likely (OR= 1.310 [95% CIs: 1.122-1.528]) to have a profile of increasing behavioural and emotional symptoms in early childhood, followed by elevated symptoms of conduct disorder, oppositionality, hyperactivity and inattention by age 8. Schizophrenia-associated alleles are linked to specific patterns of early-life psychopathology. The associations are small, but findings of this nature can help us better understand the developmental emergence of schizophrenia.

Introduction

Genetic risk for schizophrenia is highly polygenic^{1,2}, representing an additive combination of many common genetic variants with small effects. Large-scale genome-wide association studies (GWAS)³ discover these effects, with currently more than 176 genetic loci identified as conferring risk for schizophrenia⁴. Effect sizes and standard errors from GWAS can be combined across many variants in polygenic risk scores (PRS)⁵, to examine how common genetic liability for schizophrenia is associated with traits and behaviours in the general population⁶.

Schizophrenia is most commonly diagnosed in late adolescence or early adulthood⁷. PRS-based approaches can be used to study early-life manifestations of genetic risk for schizophrenia, with the potential to inform as to how and why the disorder emerges later in life. Previous studies have shown that genetic liability for schizophrenia is modestly (typically explaining <1% phenotypic variance) but robustly associated with a range of childhood outcomes, including infant neuromotor development⁸, early neurocognitive and behavioural development⁹, sleep problems¹⁰ and social cognition¹¹ – as well as with measures of psychopathology (including symptoms of anxiety, depression, attention deficit hyperactivity disorder, and conduct problems) across childhood^{12–15}. These associations appear to persist into adolescence^{16,17} as well as potentially diversifying further (e.g., into disordered eating¹⁸ and cannabis use¹⁹). The breadth of associations is underlined by findings that schizophrenia PRS explain variance in a latent, general “*p*” (for “*psychopathology*”) factor in both childhood¹⁴ and adolescence²⁰. The “*p*” factor model²¹ offers a parsimonious explanation for comorbidity and shared genetic influence between symptom domains.^{22,23}

Although it seems clear that early life manifestations of schizophrenia genetic liability across the psychopathological “phenome” are relatively diverse, it is not well understood how this diversity arises. For example, it could be that schizophrenia-associated genes have transient, developmentally-varying effects; or that early effects in specific domains trigger developmental

cascades that encompass a wider range of behaviours. It could simply be the case that these genes have highly generalized effects on behaviour; or that the environmental and maturational restrictions of childhood make their expression more diffuse than is observed in adulthood. To narrow down the various possible explanations, it is necessary to investigate the characteristics of associations in two additional dimensions: developmental time, and phenotypic space. That is, to ask to what degree effects are developmentally stable (influencing behaviour consistently across development) vs. age-specific (transient or emerging at a specific point), and to explore how broadly vs. selectively they influence different behaviours and symptoms.

In the current study, we use data from the population-based Norwegian Mother, Father and Child Cohort Study (MoBa) ²⁴ to investigate the extent to which manifestations of genetic liability for schizophrenia in childhood emotional and behavioural problems are: i) developmentally stable vs. unstable (age-specific); and ii) domain-general vs. domain-specific. Additionally, we aim to investigate genomic prediction, using schizophrenia PRS, of individuals' latent symptom profiles across development and a broad range of symptom domains. The purpose of this facet of our analysis is to explore whether combining information about symptoms of psychopathology across time and domains is a useful way to maximise the explanatory power of genomic information in a developmental context. ²⁵

The analyses we present are, in large part, exploratory. However, for the first two aims, in each case we seek to identify the most parsimonious model describing the covariation between schizophrenia PRS and various questionnaire measures of childhood emotional and behavioural problems. Based on previous findings, we had two explicit hypotheses corresponding to these aims. First, we expected developmentally-stable associations between schizophrenia-associated genetic variants and symptoms of psychopathology across early and middle-childhood, based on both previous work with schizophrenia PRS ^{12,13,15} and findings of the stable genetic influence on childhood psychopathology ²⁶. Second, we expected associations later in childhood to be primarily domain-general, mediated via a latent, general “*p*” factor.

Materials and methods

Study sample

The Norwegian Mother, Father and Child Cohort Study (MoBa) is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health ²⁷. Participants were recruited from all hospitals and obstetric units in Norway during 1999-2008. The women consented to participation in 41% of the pregnancies. The cohort now includes 114,500 children, 95,200 mothers and 75,200 fathers. The current study is based on version 12 of the quality-assured data files released in January 2019.

The establishment and data collection in MoBa were previously based on a license from the Norwegian Data Protection Agency and approval from The Regional Committee for Medical Research Ethics (REK) and is now based on regulations under the Norwegian Health Registry Act. The current study was approved by REK (2016/1702).

Blood samples were obtained from children (umbilical cord) at birth. Genotyping of the entire MoBa cohort is ongoing. We used genotype data from children in 17,000 randomly selected trios, and after quality control the analytic sample was a genotyped sub-set (N = 15,105) of children ²⁴. Details about the processing of the genetic data are outlined in the Supplementary Information.

Measures

For the longitudinal analyses across early childhood (18 months, 3 years, 5 years), we used maternal reports of children's emotional problems (5 items) and behavioural problems (10 items), on sub-scales of the Child Behaviour Checklist (CBCL).²⁸ For the specificity analyses, we used three maternal report scales measuring six distinct domains of emotional and behavioural psychopathology in middle childhood (at the 8-year data collection): the 13-item Short Mood and

Feelings Questionnaire (sMFQ) ²⁹ measured symptoms of depression, a 5-item short form of the Screen for Child Anxiety Related Disorders (SCARED) ³⁰ measured symptoms of anxiety, and a 34-item version of the Rating Scale for Disruptive Behaviour Disorders (RS-DBD ³¹) measured symptoms of conduct problems (CD), oppositional defiant disorder (ODD), hyperactivity, and inattention. Details about measure selection and psychometrics are in the Supplementary Information.

Polygenic risk scores

We calculated PRS for schizophrenia using PRSice2 ³², based on European samples from the most recent ² Psychiatric Genomics Consortium schizophrenia genome-wide association study (GWAS). PRS can be calculated using effect estimates for all variants in common between the discovery (i.e., GWAS) sample and target sample, for variants whose effects in the GWAS had a p-value below a specified threshold. Typically, PRS are created at a range of p-value thresholds (0-1), reflecting the expectation that the polygenic signal will gradually increase as more variants (either with weaker effects or a lower frequency in the population) are included, up until the point additional variants contribute only statistical noise to the score. We opted to use the thresholds from the polygenic scoring analysis presented in the original GWAS ². These were: $p < 0.001$, $p < 0.01$, $p < 0.05$, $p < 0.1$, $p < 0.2$, $p < 0.5$, $p < 1$. Full details of the parameters and procedure used in generating the PRS are in the Supplementary Information.

Analyses

Modelling developmental stability

Latent growth models were used to explore variability in the development of emotional and behavioural problems between the ages of 18 months and 5 years in the context of genetic liability for schizophrenia. In these models, variance in observed variables is explained by two latent growth

factors: a latent intercept, which is specified to load equally at all waves (here at 1.5, 3, and 5 years), and a latent slope, loadings for which are fixed proportional to their temporal distance from the first wave of measurement (here the loadings are 0, 1.5, and 3). For each schizophrenia PRS (based on the different thresholds), we formally compared the fit of models specifying an effect on growth parameters (intercept and slope) vs age-specific residuals in emotional and behavioural problems. This model fitting indicated to what extent the association between schizophrenia PRS and emotional/behavioural problems is developmentally stable (i.e., via latent growth factors), age-specific (i.e., via residuals), or null. Informal comparisons were also made to ascertain whether stable effects could be primarily ascribed to either the latent intercept or slope factor (for full details of model fitting and comparison strategy, see Supplementary Information).

Modelling phenotypic specificity

Next, we specified models describing covariation among items from the 8-year psychopathology variables with 6 domain-specific factors and one overall “ p ” factor. We compared their fit to the data when schizophrenia PRS were allowed to influence, respectively: symptom-specific residuals, domain-specific factors, the domain-general “ p ” factor, as well specified to have no effect. All versions of a given model were nested, so formal tests of fit (χ^2 difference tests) were used to ascertain the best-fitting model.

For both the developmental and “ p ” factor modelling, after performing the formal model fitting to establish the structure of the best-fitting model for each PRS threshold, we calculated the average R^2 value among the relevant outcomes for that model. In the main text, we present the results for the best-fitting model at the PRS threshold at which R^2 was maximised.

Exploring developmental symptom profiles

Finally, we incorporated both the developmental modelling and “*p*” factor modelling approaches into a latent profile analysis. In latent profile analysis, a categorical latent variable is used to assign individuals in a sample into one of a pre-specified number of profiles based on their pattern of scores across multiple observed continuous variables. For our analysis, profile membership was informed by individuals’ estimated scores on the continuous latent growth factors from the emotional and behavioural problems latent growth models, and their observed scores on compiled scale versions of the 6 psychopathology domains measured at 8 years. We then included schizophrenia PRS as predictors of profile assignment. The full model is displayed in **Figure 1**. We specified models with 2, 3, 4, 5, and 6 profiles respectively, and assessed them using a combination of standard criteria (Vuong–Lo–Mendell–Rubin likelihood ratio test^{33,34}, entropy, fit indices).

All modelling for the first two parts of the analyses was carried out in R version 3.4.4 using the *lavaan* package³⁵ version 0.6.3, and the latent profile analysis required the *Mplus* software³⁶ version 8.1 via the R package *MplusAutomation*³⁷ version 0.7.3, as *lavaan* does not support categorical latent variables at present. All models included sex as a covariate to prevent mean differences (e.g., in depressive symptoms in girls vs. boys) from having an influence on the model fit comparisons. Further details are available in the Supplementary Information, and code for the modelling is openly available at <https://github.com/psychgen/scz-prs-psychopathol-dev>. The consent given by the participants does not allow for storage of data on an individual level in repositories or journals. Researchers can apply for access to data for replication purposes via MoBa, in line with their data access policies.

Results

Descriptive statistics for all study variables are available in **Table 1**, alongside an exploration of the extent of selection bias among our genotyped sub-sample of MoBa and of selective attrition with respect to the longitudinal analyses (both of which were observed at low levels) is included in sTables 2 and 3.

Developmental stability analyses

Linear growth models provided an acceptable fit to the data from the measures of emotional and behavioural problems (for details, see Supplementary Information). We found evidence of associations between schizophrenia genetic liability and emotional and behavioural problems, with the null model rejected at most p-value thresholds in each domain. Moreover, we observed evidence in favour of developmentally stable rather than age-specific associations, with models incorporating PRS effects on the latent growth factors preferred in both domains (see [Table 2](#)). For emotional problems, formal model fit comparisons did not indicate whether these effects were primarily mediated via the intercept (i.e., predicting the overall level of symptoms) or slope (i.e., rate of change in symptoms over time) in the best-fitting model. That is, both models represented an acceptable constraint to the model with effects on both intercept and slope, and neither model could be rejected in favour of a more parsimonious model with no developmentally-stable effects. Point estimates were similar for both ($\beta_{\text{intercept}} = 0.019$, 95% CI: -0.005–0.043; $\beta_{\text{slope}} = 0.019$, 95% CI: -0.010–0.048; PRS threshold: $p < 0.1$). In contrast, for behavioural problems, these effects were mediated via the slope factor alone at the most predictive threshold ($\beta_{\text{slope}} = 0.032$; 95% CI 0.007–0.057; threshold: $p < 1$), meaning that early symptoms of behavioural problems were more persistent across early childhood as schizophrenia PRS increased. However, in both cases the associations were small, explaining ~0.1% variance in the growth factors. Fit statistics and parameter estimates for all models are included in the Supplementary Information.

Phenotypic specificity analyses

A structural model of domain-specific factors and one overarching “*p*” factor provided a good fit to item-level data from the 8-year psychopathology measures (CFI=0.98, TLI=0.97, RMSEA=0.04). Details of the model fitting process are presented in the Supplementary Information. We found that schizophrenia PRS were significant predictors of variability in symptoms of psychopathology at 8

years, and that this prediction was maximised at the PRS threshold including all variants ($p < 1$). At this threshold (as well as three others tested: $p < 0.05$, $p < 0.2$, and $p < 0.5$), neither the model incorporating a PRS effect on the general “ p ” factor nor the model incorporating domain-specific effects fit the data sufficiently well to be accepted. Instead, the preferred model allowed the PRS to influence item-specific residuals – i.e., explaining symptom-specific variation. **Figure 2** shows standardized beta coefficients for these symptom-level relationships.

The pattern of results in **Figure 2** show that there is substantial heterogeneity, in terms of schizophrenia PRS associations, between symptoms – even within domains. This is particularly evident for symptoms within the depression and inattention domains. Nonetheless, as with the developmental modelling, effect sizes were very small, with PRS explaining a maximum of 0.4% variance in symptoms. Furthermore, at other PRS thresholds, both the PRS-on-“ p ” factor ($p < 0.001$ and $p < 0.01$) and PRS-on-domain-specific factors model ($p < 0.01$) were accepted. Parameter estimates and 95% CIs for these alternatives are presented in the Supplementary Information.

Developmental symptom profiles

Scores from all scales (CBCL measures of emotional and behavioural problems at 18 months, 3 years, and 5 years and six 8-year psychopathology domains) were incorporated into a single model for the latent profile analysis. Fit statistics for these models are presented in the Supplementary Information. Entropy (reflecting the overall certainty with which individuals could be assigned to symptom profiles) was good (~ 0.8) for each version of the model, but the Vuong–Lo–Mendell–Rubin test statistic for the 5-profile model indicated that it offered no significant improvement of fit on the simpler 4-profile model.

Figure 3 shows the 4 symptom profiles. Individuals assigned to profile 1 (5.7% of sample) had increasing symptoms of emotional problems between 18 months and 5 years and elevated anxiety symptoms at age 8. Profile 2 (7.9%) was characterised by moderate, stable behavioural problems

symptomatology in early childhood, and moderate symptoms of ODD, hyperactivity, and inattention at 8 years. Profile 3 was considered the normative symptom profile, as a large majority (84.7%) could be assigned to it, and it was characterised by decreasing problems developmentally and low levels of symptoms at 8 years. Profile 4 was the least populous (1.8% of the sample) and most symptomatic profile, characterised by increasing/elevated symptoms of behavioural problems and increasing symptoms of emotional problems preceding relatively elevated symptoms of conduct disorder, ODD, hyperactivity and inattention at 8 years.

The probabilities of classification into profiles 1, 2, and 4 relative to the normative profile are shown in [Figure 4](#) as a function of PRS (at the $p < 0.05$ threshold), alongside odds ratios of PRS on probability of classification into a given profile rather than the normative profile. Individuals with higher PRS were more likely to be assigned to profile 4 than the normative profile (OR= 1.310 [95% CIs: 1.122-1.528], bottom-right of figure). Profile 4 was also better defined – relative to the normative profile – than profiles 1 or 2: this is evident, in the bottom-left panel of [Figure 4](#), in the density of high probabilities of assignment into this profile, illustrated by the proximity of the contour lines at the top of the panel, and in the increased separation between the different coloured contour lines in this panel (indicating that very few individuals were equally likely to be assigned to either Profile 4 or the normative profile). There was no strong evidence of associations between schizophrenia PRS and probability of assignment to either of the other profiles (relative to the normative profile).

Discussion

The results indicate that associations between schizophrenia risk-associated alleles and symptoms of early childhood (18 months to 5 years) psychopathology tend to be developmentally stable, rather than transient or emergent. In middle childhood (8 years) we found that, contrary to expectations, associations between schizophrenia genetic risk and symptoms of psychopathology may be better explained by symptom-specific effects than by prediction of a latent “ p ” factor indexing general

psychopathology. Finally, we showed that by combining information about symptoms of psychopathology across time and domains, associations with schizophrenia genetic liability can be summarised in the form of a characteristic symptom profile.

Our results are supportive of the neurodevelopmental model of schizophrenia³⁸ and consistent with previous findings of widespread associations between schizophrenia genetic liability and symptoms of childhood psychopathology. There are a number of possible explanations for such associations. They may arise because of *pleiotropy*, that is, direct effects of schizophrenia-related genetic variants on other behaviours and symptoms. They may represent early manifestations of schizophrenia liability, where environmental or maturational restrictions mean that the classic “schizophrenia phenotype” (and its sub-clinical analogues) cannot be fully expressed until adolescence or later. A further alternative is that (some) childhood manifestations of genetic liability for schizophrenia are actually part of a causal process in the development of the disorder. This could involve intrinsic, developmental processes (e.g., specific thought patterns becoming “grooved” over time) or extrinsic causal processes, where genetic risk is mediated via the environment. An example of such a process would be the putatively causal link between cannabis use and schizophrenia^{39,40}. Cannabis is an environmental exposure, but its use is associated with genetic variants linked to impulsivity⁴¹ and risk-taking⁴². Assuming a demonstrable causal relationship between cannabis use and the development of schizophrenia, the environment (cannabis use) would *mediate* the effects of these genetic variants on schizophrenia; and impulsivity and risk-taking earlier in development would be on the causal pathway.

We applied competing structural models to associations between schizophrenia PRS and symptoms of childhood psychopathology in order to go beyond simply detecting or even quantifying these links, seeking instead to explore their characteristics across developmental time and phenotypic space. For example, using developmental models, we found that associations between schizophrenia PRS and emotional and behavioural symptomatology were best explained via effects on developmentally stable growth processes. This is in line with evidence on the stability of genetic influences, in general, across childhood²⁶. We had limited power to investigate the nature of these

stable associations, but found evidence that, for behavioural problems, schizophrenia-associated alleles are primarily predictive of individual differences in their propensity to change over time in early childhood. Such a pattern is more consistent with an aetiological model implicating environmentally mediated genetic effects and active or evocative gene-environment correlations, than direct pleiotropy. Future work incorporating measured environmental mediators may be able to establish whether these mechanisms are involved in the childhood manifestation of schizophrenia genetic liability.

The “ p ” factor structure of childhood psychopathology has been largely supported by previous genetically-informative work indicating that genetic prediction of symptoms is primarily mediated via a domain-general factor (e.g., ²²). Schizophrenia genetic risk has been shown to associate with psychopathology via this route ^{14,16}. In apparent contrast, we found that models with symptom-specific effects were preferred, despite the cost of estimating many more parameters. There is likely some overfitting involved in these models, which is why we do not interpret the individual symptom-level effects. At the domain level, schizophrenia PRS were most predictive of behavioural problems such as conduct and oppositional defiance. This finding contrasts somewhat with prior evidence of marginally stronger effects for emotional problems such as anxiety and depression, albeit primarily in older children and teenagers.¹² It is possible that the more behavioural manifestation we observed is somewhat childhood-limited. Alternatively, it could be that measurement differences account for these contrasting patterns. What is clear is that there is substantial heterogeneity of these effects – even within domains. This strengthens the conclusion that genetic liability for schizophrenia influences childhood psychopathology in a more specific and less generalized manner than has been shown previously.^{12,13,15}

It is likely that two aspects of our approach explain why we uncover this specificity. First, we used item-level data, providing more scope for heterogeneity – from both signal and noise – to emerge, whereas scale-level data has often been used in the past. Second, we tested schizophrenia PRS at different thresholds – including thresholds at which all or almost all SNPs are included in the score. Previous work has often selected a single threshold score, typically $p < 0.05$ or lower. We found that

lower threshold scores were likely to have smaller associations that, partly for reasons of statistical power, led to more parsimonious models being preferred. However, the amount of variance explained in symptoms continued to increase for scores created with less stringent thresholds suggesting that a single-score approach to using PRS potentially leads to meaningful signal from SNPs being missed. Using scores at higher thresholds both increases power and is consistent with the theory underpinning the polygenic model of genetic effects ⁵.

The small effect sizes for schizophrenia PRS on measures of childhood psychopathology in our sample are consistent with the literature ^{9,12,15}. Our latent profile analytic approach demonstrates one way in which combining information across time and symptom domains can draw out more meaningful effect sizes (i.e., we observed effects equivalent to a 30% increase in the probability of displaying a particular symptom profile per standard deviation increase in schizophrenia PRS). However, we acknowledge the inherent bounds on the utility of effects explaining so little variance in childhood and adolescent outcomes. Using PRS for adult psychiatric disorders to identify children and adolescents at substantially increased, personal risk for behavioural and emotional difficulties and basing targeted prevention strategies on this information is currently not a realistic prospect. That is, with effects of the magnitude observed in these analyses, schizophrenia PRS have little or no clinical utility at the individual-level of in a paediatric population. As increasing GWAS sample sizes and methodological improvements allow polygenic prediction to increase and account for more so-called “missing heritability” ⁴³, it is possible that this situation may evolve. Nonetheless, even with the current barriers to clinical utility, we show here that PRS do have value as tools to help clarify the mechanisms by which symptoms emerge, are maintained, differentiate, exacerbate or improve developmentally.

Our study is subject to several limitations. Small effect sizes for schizophrenia PRS on childhood traits mean that, even with a sample of several thousand children, we have limited power for differentiating similar parameterisations (e.g., on latent growth processes) and therefore opted against stratifying by sex. Secondly, the genotyped sample we use is subject to some selection effects, and the MoBa sample as a whole is (like all cohort studies) affected by selective attrition as

the propensity to drop out of studies over time is linked to poorer health, which is also reflected genetically⁴⁴. Furthermore, the measures used to index symptoms of emotional and behavioural psychopathology are relatively brief and available only as maternal reports. Some, most notably both CBCL sub-scales at 18 months, 3 years, and 5 years and the anxiety measure at 8 years, have lower internal consistency than is desirable, and statistical noise in these measures will decrease power to detect associations with polygenic scores. Corroboration by additional reporters and using clinical interviews or diagnoses from linked health care registries would help to mitigate this limitation in future work.

Overall, our findings suggest that schizophrenia-related genetic variants are associated with symptoms of psychopathology stably from as early as 18 months of age, and with considerable phenotypic specificity by middle childhood. Further work – such as expanding models longitudinally and to incorporate measured environments and testing within-family polygenic prediction – can help to refine and triangulate developmental explanations for the emergence of schizophrenia later in life.

Tables

Measure	Age (yrs)	N	Mean	SD	Min	Max	Skew	Kurtosis	α
Behavioural problems (CBCL)	1.5	11552	3.87	2.21	0	16	0.63	0.42	0.70
	3	9369	3.78	2.38	0	16	0.75	0.47	0.77
	5	7274	2.42	2.23	0	15	1.21	1.69	0.81
Emotional problems (CBCL)	1.5	11549	1.26	1.18	0	8	1.01	1.11	0.66
	3	9370	1.33	1.35	0	9	1.24	1.89	0.69
Depressive symptoms (sMFQ)	5	7272	1.03	1.26	0	10	1.64	3.59	0.73
	8	7111	1.78	2.38	0	23	2.36	8.40	0.92
Anxiety symptoms (SCARED)	8	7118	1.02	1.20	0	10	1.67	4.27	0.76
Conduct disorder symptoms (RS-DBD)	8	7121	0.76	1.50	0	16	3.13	13.88	0.88
Oppositional defiant disorder symptoms (RS-DBD)	8	7111	3.41	3.11	0	22	1.45	3.39	0.91
Hyperactivity symptoms (RS-DBD)	8	7114	3.47	3.87	0	27	1.91	4.96	0.91
Inattention symptoms (RS-DBD)	8	7115	4.82	4.01	0	27	1.67	4.07	0.92

Table 1. Descriptive statistics for main study variables (scale-level only)

Note: α = ordinal Cronbach's alpha; CBCL = Child Behavior Checklist; sMFQ = Short Mood and Feelings Questionnaire; SCARED = Screen for Child Anxiety Related Disorders; RS-DBD = Rating Scale for Disruptive Behaviour Disorders;

Scale	PRS Threshold*	Model (Comparator)	Df	AIC	Δ Chisq	p
Behavioural problems	p<1	Age-specific PRS effects (none)	1	122345.23		
		PRS effects on latent growth – both (age-specific)	2	122343.62	0.386	0.534
		PRS effects on latent intercept only (age-specific)	3	122346.56	4.945	0.026
		No PRS effects (intercept only)	4	122345.98	1.419	0.234
		PRS effects on latent slope only (age-specific)	3	122341.63	0.018	0.892
		No PRS effects (slope only)	4	122345.98	6.345	0.012
		No PRS effects (latent growth - both)	4	122345.98	6.364	0.042
Emotional problems	p<0.1	Age-specific PRS effects (none)	1	90224.89		
		PRS effects on latent growth – both (age-specific)	2	90223.27	0.379	0.538
		PRS effects on latent intercept only (age-specific)	3	90222.95	1.683	0.195
		No PRS effects (intercept only)	4	90227.72	6.764	0.009
		PRS effects on latent slope only (age-specific)	3	90223.64	2.374	0.123
		No PRS effects (slope only)	4	90227.72	6.073	0.014
		No PRS effects (latent growth - both)	4	90227.72	8.447	0.015

Table 2. Genetic model-fitting results from the developmental analyses

Notes: *denotes PRS threshold at which explained variance was maximized – for all model-fitting results from all tested thresholds see sTable 6; Df= model degrees of freedom; AIC = Akaike’s Information Criterion; Δ Chisq = change in chi square comparing nested models; bold typeface shows best-fitting model(s)

Author contributions

LJH and AH conceived of the analyses; LJH, RBA, HA, and AH refined the analysis plan, prepared the polygenic scores and conducted the analyses; LJH, EC, and ØH performed quality control of the genetic data; PMM, PN, OAA, and TRK carried out the genotyping; LJH and AH drafted the manuscript; all authors critically reviewed and commented upon the manuscript.

Acknowledgements

We are grateful to all the participating families in Norway who take part in this on-going cohort study.

The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research.

We thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of Norway (NRC) (#229624). We also thank the NORMENT Centre for providing genotype data, funded by NRC (#223273), South East Norway Health Authority and KG Jebsen Stiftelsen. Further we thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data.

Funding

Laurie Hannigan, Ziada Ayorech, and Alexandra Havdahl were supported by grants from the South-Eastern Norway Regional Health Authority (2018059, 2019097 and 2018058, respectively). Ragna Bugge Askeland and Elizabeth Corfield were supported by a grant from the Norwegian Research Council (274611). Ted Reichborn-Kjennerud was supported by a grant from the Research Council of Norway (274611). Pål Rasmus Njølstad was funded by the ERC AdG SELECTIONPREDISPOSED

(#293574), the Stiftelsen Kristian Gerhard Jebsen, the Trond Mohn Foundation, the Norwegian Research Council (#240413/F20), the Novo Nordisk Foundation (#54741), the University of Bergen, and the Western Norway health Authorities (Helse Vest; PERSON-MED-DIA and #911745). Ole A. Andreassen was supported the Research Council of Norway (229129; 213837; 248778; 223273; 249711); the South-East Norway Regional Health Authority (2017-112); KG Jebsen Stiftelsen (SKGJ) and H2020 grant # 847776 CoMorMent. George Davey Smith works in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_00011/1). This work was partly supported by the Research Council of Norway through its Centre of Excellence funding scheme, project number 262700.

Role of the funding sources

The funders had no role in the direction or content of the manuscript.

Conflicts of interest

None.

References

1. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci.* 2014;17(6):782-790. doi:10.1038/nn.3708
2. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S, Neale BM, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-427. doi:10.1038/nature13595
3. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet.* 2005;6(2):95-108. doi:10.1038/nrg1521
4. Lam M, Chen CY, Li Z, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet.* 2019;51(12):1670-1678. doi:10.1038/s41588-019-0512-x
5. Dudbridge F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genet.* Published online 2013. doi:10.1371/journal.pgen.1003348
6. Mistry S, Harrison JR, Smith DJ, Escott-Price V, Zammit S. The use of polygenic risk scores to identify phenotypes associated with genetic risk of schizophrenia: Systematic review. *Schizophr Res.* 2018;197(November 2017):2-8. doi:10.1016/j.schres.2017.10.037
7. Kessler RC, Angermeyer M, Anthony JC, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization ' s. *World Psychiatry* 2007;6:168-176). 2007;6(October):168-176. doi:10.1001/archpsyc.62.6.593
8. Serdarevic F, Jansen PR, Ghassabian A, et al. Association of Genetic Risk for Schizophrenia and Bipolar Disorder With Infant Neuromotor Development. *JAMA Psychiatry.* 2018;75(1):96. doi:10.1001/jamapsychiatry.2017.3459
9. Riglin L, Collishaw S, Richards A, et al. Schizophrenia risk alleles and neurodevelopmental outcomes in childhood: a population-based cohort study. *The Lancet Psychiatry.* 2017;4(1):57-62. doi:10.1016/S2215-0366(16)30406-0
10. Reed ZE, Jones HJ, Hemani G, Zammit S, Davis OSP. Schizophrenia liability shares

common molecular genetic risk factors with sleep duration and nightmares in childhood [version 2; peer review: 2 approved]. *Wellcome Open Res.* 2019;4:1-25.

doi:10.12688/wellcomeopenres.15060.2

11. Germine L, Robinson EB, Smoller JW, et al. Association between polygenic risk for schizophrenia, neurocognition and social cognition across development. *Transl Psychiatry.* 2016;6(10):e924. doi:10.1038/tp.2016.147
12. Nivard MG, Gage SH, Hottenga JJ, et al. Genetic Overlap Between Schizophrenia and Developmental Psychopathology: Longitudinal and Multivariate Polygenic Risk Prediction of Common Psychiatric Traits During Development. *Schizophr Bull.* 2017;43(6):1197-1207. doi:10.1093/schbul/sbx031
13. Riglin L, Collishaw S, Richards A, ... AKT-P, undefined 2017. The impact of schizophrenia and mood disorder risk alleles on emotional problems: investigating change from childhood to middle age. *Psychol Med.* Published online 2017.
14. Riglin L, Thapar AKA, Leppert B, et al. Using Genetics to Examine a General Liability to Childhood Psychopathology. *Behav Genet.* 2020;50(0123456789):213-220. doi:10.1007/s10519-019-09985-4
15. Jansen PR, Polderman TJC, Bolhuis K, et al. Polygenic scores for schizophrenia and educational attainment are associated with behavioural problems in early childhood in the general population. *J Child Psychol Psychiatry Allied Discip.* 2018;59(1):39-47. doi:10.1111/jcpp.12759
16. Jones HJ, Heron J, Hammerton G, et al. Investigating the genetic architecture of general and specific psychopathology in adolescence. *Transl Psychiatry.* 2018;8(1):145. doi:10.1038/s41398-018-0204-9
17. 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(3):221-228. doi:10.1001/jamapsychiatry.2015.3058
18. Solmi F, Mascarell MC, Zammit S, Kirkbride JB, Lewis G. Polygenic risk for schizophrenia, disordered eating behaviours and body mass index in adolescents. *Br J Psychiatry.*

- 2019;215(1):428-433. doi:10.1192/bjp.2019.39
19. Hiemstra M, Nelemans SA, Branje S, et al. Genetic vulnerability to schizophrenia is associated with cannabis use patterns during adolescence. *Drug Alcohol Depend.* Published online 2018. doi:10.1016/j.drugalcdep.2018.05.024
 20. Leppert B, Havdahl A, Riglin L, et al. Association of maternal neurodevelopmental risk alleles with early-life exposures. *JAMA Psychiatry.* 2019;76(8):834-842. doi:10.1001/jamapsychiatry.2019.0774
 21. Caspi A, Houts RM, Belsky DW, et al. The p Factor. *Clin Psychol Sci.* 2014;2(2):119-137. doi:10.1177/2167702613497473
 22. Selzam S, Coleman JRI, Caspi A, Moffitt TE, Plomin R. A polygenic p factor for major psychiatric disorders. *Transl Psychiatry.* 2018;8(1):205. doi:10.1038/s41398-018-0217-4
 23. Allegrini AG, Cheesman R, Rimfeld K, et al. The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *bioRxiv.* Published online 2019:591354. doi:10.1101/591354
 24. Magnus P, Birke C, Vejrup K, et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol.* 2016;45(2):382-388. doi:10.1093/ije/dyw029
 25. Lubke GH, Miller PJ, Verhulst B, et al. A powerful phenotype for gene-finding studies derived from trajectory analyses of symptoms of anxiety and depression between age seven and 18. *Am J Med Genet Part B Neuropsychiatr Genet.* 2016;171(7):948-957. doi:10.1002/ajmg.b.32375
 26. Hannigan LJ, Walaker N, Waszczuk MA, McAdams TA, Eley TC. Aetiological Influences on Stability and Change in Emotional and Behavioural Problems across Development: A Systematic Review. *Psychopathol Rev.* 2017;a4(1):52-108. doi:10.5127/pr.038315
 27. Magnus P, Irgens LM, Haug K, et al. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol.* 2006;35(5):1146-1150. doi:10.1093/ije/dyl170
 28. Achenbach TM, Ruffle TM. The Child Behavior Checklist and Related Forms for Assessing Behavioral/Emotional Problems and Competencies. *Pediatr Rev.* 2007;21(8):265-271. doi:10.1542/pir.21-8-265

29. Angold A, Costello EJ, Messer SC, Pickles A, Winder F, Silver D. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Methods Psychiatr Res.* 1995;5:237-249.
30. Birmaher B, Khetarpal S, Brent D, et al. The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *J Am Acad Child Adolesc Psychiatry.* 1997;36(4):545-553. doi:10.1097/00004583-199704000-00018
31. Silva RR, Alpert M, Pouget E, et al. A rating scale for disruptive behavior disorders, based on the DSM-IV item pool. *Psychiatr Q.* 2005;76(4 SPEC. ISS.):327-339. doi:10.1007/s11126-005-4966-x
32. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics.* Published online 2015. doi:10.1093/bioinformatics/btu848
33. Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. *Biometrika.* Published online 2001. doi:10.1093/biomet/88.3.767
34. Vuong QH. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica.* Published online 1989. doi:10.2307/1912557
35. Rosseel Y. lavaan : An R Package for Structural Equation Modeling . *J Stat Softw.* Published online 2015. doi:10.18637/jss.v048.i02
36. Muthén LK, Muthén BO. *Mplus User's Guide.* Eighth Edi. Muthén & Muthén
37. Hallquist MN, Wiley JF. MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Struct Equ Model.* Published online 2018. doi:10.1080/10705511.2017.1402334
38. Rapoport JL, Giedd JN, Gogtay N. Neurodevelopmental model of schizophrenia: Update 2012. *Mol Psychiatry.* 2012;17(12):1228-1238. doi:10.1038/mp.2012.23
39. Vaucher J, Keating BJ, Lasserre AM, et al. Cannabis use and risk of schizophrenia: A Mendelian randomization study. *Mol Psychiatry.* 2018;23(5):1287-1292. doi:10.1038/mp.2016.252
40. Gage SH, Jones HJ, Burgess S, et al. Assessing causality in associations between cannabis use and schizophrenia risk: A two-sample Mendelian randomization study. *Psychol Med.*

2017;47(5):971-980. doi:10.1017/S0033291716003172

41. Soler Artigas M, Sánchez-Mora C, Rovira P, et al. Attention-deficit/hyperactivity disorder and lifetime cannabis use: genetic overlap and causality. *Mol Psychiatry*. Published online 2019. doi:10.1038/s41380-018-0339-3
42. Strawbridge RJ, Ward J, Cullen B, et al. Genome-wide analysis of self-reported risk-taking behaviour and cross-disorder genetic correlations in the UK Biobank cohort. *Transl Psychiatry*. 2018;8(1):1-11. doi:10.1038/s41398-017-0079-1
43. Young AI. Solving the missing heritability problem. *PLoS Genet*. 2019;15(6):1-7. doi:10.1371/journal.pgen.1008222
44. Adams MJ, Hill WD, Howard DM, et al. Factors associated with sharing e-mail information and mental health survey participation in large population cohorts. *Int J Epidemiol*. Published online July 1, 2019. doi:10.1093/ije/dyz134

Figure 1. Latent profile analysis to ascertain developmental and domain-specific profiles associated of psychopathology symptoms associated with schizophrenia genetic effects

Note- boxes represent observed variables and circles model-estimated latent variables; I = intercept factor, which loads equally on observed variables at all waves; C = categorical latent variable, sub-dividing the sample into a specified number of classes (we tested models with 2, 3, 4, 5, and 6) according to values on: S = slope factor with loadings 0, 1.5, 3.5, corresponding to temporal distance from the first wave of measurement; EMO = CBCL emotional problems symptoms; BEH = CBCL behavioural problems symptoms; DEP = sMFQ depressive symptoms; ANX = SCARED anxiety symptoms; CD = RS-DBD conduct disorder symptoms; ODD = RS-DBD oppositional defiant disorder symptoms; HYP = RS-DBD hyperactivity (ADHD) symptoms; ODD = RS-DBD inattention (ADHD) symptoms; 8-year observed variables and emotional/behavioural intercept/slope variables, respectively, are inter-correlated within class (paths omitted from diagram for clarity)

Figure 2. Item-level heterogeneity of schizophrenia PRS associations with 8-year childhood psychopathology

Note- Bars indicate 95% confidence intervals; PRS threshold $p < 1$ (i.e., all variants)

Figure 3. Symptom profiles from latent profile analysis of symptoms of psychopathology across development and domain

Note- Bars/bands indicate 95% confidence intervals; %s in header indicate the proportion of the sample best classified in each profile; DEP = sMFQ depressive symptoms; ANX = SCARED anxiety symptoms; CD = RS-DBD conduct disorder symptoms; ODD = RS-DBD oppositional defiant disorder symptoms; HYP = RS-DBD hyperactivity (ADHD) symptoms; ODD = RS-DBD inattention (ADHD) symptoms; 8-year observed variables and internalizing/externalizing intercept/slope variables, respectively, are inter-correlated within class (paths omitted from diagram for clarity)

Figure 4. Density of relative probabilities of classification into the different profiles by schizophrenia PRS and odds ratios for prediction of symptom profile classification by PRS

Note- contour plots show probabilities of classification into a specific profile (1, 2, or 4) relative to probabilities of classification into the normative profile (3) for individuals ultimately assigned to each of these profiles (indicated by colour coding), as a function of PRS; diamond marker indicates within-profile means (both classification probability and PRS) and the shaded region around them shows the 95% CIs in either dimension (i.e., vertically for probability, horizontally for PRS); profile 3 is used as reference category for logistic regression; PRS threshold used is $p < 0.05$, selected as it maximizes the OR for profile 4 vs. profile 3; but the pattern of results is consistent at other thresholds (see Supplementary Information)