




OPEN

Greater genetic risk for adult psychiatric diseases increases vulnerability to adverse outcome after preterm birth

Harriet Cullen^{1,2}, Saskia Selzam³, Konstantina Dimitrakopoulou⁴, Robert Plomin³ & A. David Edwards¹

Preterm birth is an extreme environmental stress associated with an increased risk of later cognitive dysfunction and mental health problems. However, the extent to which preterm birth is modulated by genetic variation remains largely unclear. Here, we test for an interaction effect between psychiatric polygenic risk and gestational age at birth on cognition at age four. Our sample comprises 4934 unrelated individuals (2066 individuals born < 37 weeks, 918 born < = 34 weeks). Genome-wide polygenic scores (GPS's) were calculated for each individual for five different psychiatric pathologies: Schizophrenia, Bipolar Disorder, Major Depressive Disorder, Attention Deficit Hyperactivity Disorder and Autism Spectrum Disorder. Linear regression modelling was used to estimate the interaction effect between psychiatric GPS and gestational age at birth (GA) on cognitive outcome for the five psychiatric disorders. We found a significant interaction effect between Schizophrenia GPS and GA ($\beta = 0.038 \pm 0.013$, $p = 6.85 \times 10^{-3}$) and Bipolar Disorder GPS and GA ($\beta = 0.038 \pm 0.014$, $p = 6.61 \times 10^{-3}$) on cognitive outcome. Individuals with greater genetic risk for Schizophrenia or Bipolar Disorder are more vulnerable to the adverse effects of birth at early gestational age on brain development, as assessed by cognition at age four. Better understanding of gene-environment interactions will inform more effective risk-reducing interventions for this vulnerable population.

Preterm birth is an extreme perinatal stress and is associated with a significantly increased risk of later cognitive dysfunction and psychiatric disease. Preterm birth is common, accounting for around 11% of all births and is a leading cause of infant mortality and morbidity worldwide¹. Preterm infants have higher rates of cognitive impairment, cerebral palsy, autism spectrum disorders and psychiatric disease²⁻⁵, not all of which can be explained by the degree of prematurity and severity of clinical problems around the time of birth.

Magnetic resonance imaging studies indicate that preterm brain injury is a multifactorial disorder which encompasses alterations to cortical and subcortical gray matter and white matter. These studies have contributed to our understanding of the pathophysiology underlying neurodevelopmental disability and cognitive impairments in the preterm population, however, a great deal remains unknown. Specifically, the extent to which genetic factors play a role in the variable cognitive outcomes of preterm infants and the possible interaction between genetic risk and the environmental exposure to preterm birth has not been examined in detail. Here, we explore whether common genetic variation is important in determining an individual's susceptibility to the environmental stress of prematurity and thus plays a role in the diverse neurodevelopmental and cognitive outcomes we observe in preterm infants.

Research into the genetics of psychiatric disease has shown it to be polygenic; psychiatric disorders are influenced by many genetic variants, each of small effect⁶. We therefore anticipate that gene-environment interactions involving these complex pathologies will be most effectively studied by looking at multiple genetic variants rather than focussing on a single genetic locus. As samples sizes have grown, genome-wide association studies (GWAS) have become increasingly informative, allowing detection of small effects of single nucleotide polymorphisms

¹Centre for the Developing Brain, School of Biomedical Engineering and Imaging Sciences, King's College London, London SE1 7EH, UK. ²Department of Medical and Molecular Genetics, School of Basic and Medical Biosciences, King's College London, London SE1 9RT, UK. ³Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE5 8AF, UK. ⁴Translational Bioinformatics Platform, NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust and King's College London, London SE1 9RT, UK. ✉email: harriet.cullen@kcl.ac.uk

Psychiatric GPS	GPS _{psych} × GA		
	β_3	SE	p_3
ADHD	- 0.015	0.016	0.302
ASD	- 0.003	0.015	0.846
Bipolar Disorder	0.038	0.014	6.61 × 10⁻³
Major Depressive Disorder	0.001	0.014	0.958
Schizophrenia	0.038	0.013	6.85 × 10⁻³

Table 1. Gene-environment interaction between genetic risk for psychiatric disease and gestational age at birth on cognition at four. Results exploring the effect of the interaction term GPS_{psych} × GA on cognitive outcome in 4934 unrelated individuals. β_3 is the estimate of the effect size of the interaction term GPS_{psych} × GA, SE is the standard error of the effect size and p_3 is the corresponding p-value. Bonferroni correction was used to adjust for multiple testing and p-values < 0.01 (0.05/5) were considered statistically significant. Significant results are indicated in bold. Beta regression coefficients and associated p-values for all terms in the regression model are detailed in Supplementary Table S3 in the Supplementary Information.

(SNPs). Using the summary statistics from these studies it is possible to generate individual-specific genotypic scores to predict phenotypic variance, genome-wide polygenic scores (GPS). A genome-wide polygenic score, which is an aggregate of trait-related effect sizes of SNPs across the genome in independent samples can be used to test the predictive power of multiple genetic variants simultaneously. Genome-wide polygenic scores provide a useful approach to exploring gene-environment interactions in complex traits.

In previous work, we hypothesised that genes associated with psychiatric disease might increase vulnerability to abnormal brain development in infants subjected to the environmental stress of prematurity⁷. Under this notion, individuals with a greater polygenic risk for psychiatric disease, when born preterm, would be more likely to suffer adverse developmental consequences. We showed that in a cohort of preterm infants, psychiatric GPS was negatively associated with lentiform volume; genetic predictors of neuropsychiatric disease increase vulnerability to abnormal lentiform development after perinatal stress.

We now extend this work by looking at the possible influence of genome-wide polygenic risk for psychiatric disease in preterm individuals, as measured by their cognitive abilities in early childhood. In a large cohort of unrelated twins, we test for a possible gene-environment interaction between the GPSs for five different psychiatric disorders (Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), Bipolar Disorder, Major Depressive Disorder and Schizophrenia) and gestational age at birth on cognition at age four.

Results

Linear regression modelling was used to test for a possible gene-environment interaction between the GPSs for five different psychiatric disorders (ASD, ADHD, Bipolar Disorder, Major Depressive Disorder and Schizophrenia) and gestational age at birth on cognition at age four. Genome-wide polygenic scores were adjusted for the first five ancestry principle components, genotype chip and plate and each regression model included the covariates sex and socio-economic status (SES) (Eq. (1)).

Genotype, phenotype and covariate information was available for 4934 individuals (2594 females and 2340 males). Of these 2066 were born before 37 weeks completed gestation and 918 were born at or below 34 weeks completed gestation. A table showing the number of individuals at different gestational age cut-offs is included in the Supplementary Information (Supplementary Table S2).

Results indicated a significant interaction between psychiatric genetic risk and gestational age at birth for two of the five pathologies investigated (Table 1) after Bonferroni adjustment for multiple testing ($p < 0.05/5 = 0.01$). There was evidence for an interaction effect for gestational age at birth and Schizophrenia genetic risk ($\beta_3 = 0.038$, $se = 0.013$, $p_3 = 6.85 \times 10^{-3}$) and an interaction effect for gestational age at birth and Bipolar genetic risk ($\beta_3 = 0.038$, $se = 0.014$, $p_3 = 6.61 \times 10^{-3}$). There was no evidence for an interaction effect between gestational age at birth and genetic risk for Autism Spectrum Disorder, ADHD or Major Depressive Disorder. For both Schizophrenia and Bipolar Disorder, the effect of gestational age at birth on cognitive outcome was greater for those individuals with a larger psychiatric genetic risk burden. Results for the full regression model for all five psychiatric pathologies are given in Supplementary Table S3.

Removing the covariate socio-economic status from the model did not markedly alter the results. There remained a significant interaction effect of gestational age at birth and Schizophrenia genetic risk ($\beta = 0.038$, $se = 0.013$, $p = 7.07 \times 10^{-3}$) and gestational age at birth and Bipolar genetic risk ($\beta = 0.040$, $se = 0.014$, $p = 4.60 \times 10^{-3}$) on cognitive outcome at four. Further, there was no evidence for an association of cognition with twin birth order and the addition of this as a covariate to our model did not significantly alter the results. There remained a significant interaction effect of gestational age at birth and Schizophrenia genetic risk ($\beta = 0.038$, $se = 0.014$, $p = 6.75 \times 10^{-3}$) and gestational age at birth and Bipolar genetic risk ($\beta = 0.039$, $se = 0.014$, $p = 6.00 \times 10^{-3}$) on cognitive outcome at four.

In addition to our linear regression analysis, considering gestational age and psychiatric genetic risk as continuous variables, we compared the effect of gestational age at birth on cognition for those individuals in the highest and lowest quintiles of the Schizophrenia and Bipolar Disorder GPS distributions. For this analysis we

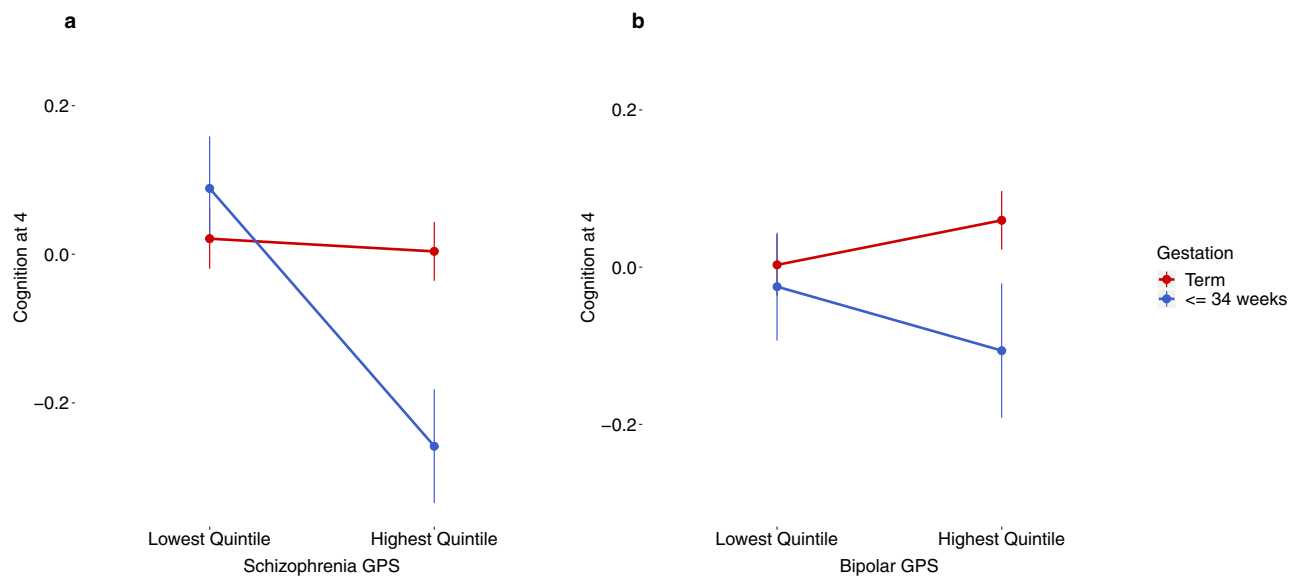


Figure 1. (a) Mean cognitive score at age four corrected for sex and SES by Schizophrenia GPS and gestational age for individuals falling in the highest and lowest 20% of the distribution of Schizophrenia GPS. There was evidence for an interaction effect ($F(1506) = 8.03$, $p = 4.65 \times 10^{-3}$); (b) Mean cognitive score at age 4 corrected for sex and SES by Bipolar disorder GPS and gestational age for individuals falling in the highest and lowest 20% of the distribution of the Bipolar Disorder GPS. No statistically significant interaction effect was found ($F(1530) = 1.55$, $P = 0.213$), GPS: genome-wide polygenic score.

selected only those individuals born at 34 weeks gestation or less and compared them with term-born individuals ($> = 37$ weeks).

Figure 1a compares the mean cognitive scores at age four corrected for sex and socio-economic status for both term-born individuals and individuals born at ≤ 34 weeks in the highest and lowest quintiles of the Schizophrenia GPS distribution. Two-way ANOVA analysis confirms a statistically significant interaction between gestational age at birth and level of Schizophrenia risk on cognition ($F(1506) = 8.03$, $p = 4.65 \times 10^{-3}$). The adverse effect of birth at earlier gestational age (≤ 34 weeks) on cognition is greater for individuals with a higher genetic risk burden for Schizophrenia than it is for those with the lowest risk (Fig. 1a and Supplementary Table S4).

Figure 1b compares the mean cognitive scores at age four, corrected for both sex and socio-economic status for both term-born individuals and individuals born at 34 weeks gestation or less in the highest and lowest quintiles of the Bipolar Disorder GPS distribution. In this case two-way ANOVA analysis did not indicate a significant interaction effect between gestational age at birth and level of Bipolar Disorder genetic risk on cognition ($F(1530) = 1.55$, $p = 0.213$). Although differences are observed, unlike Schizophrenia, the adverse effect of birth at early gestational age (≤ 34 weeks) on cognition was not significantly different when comparing those individuals at highest and lowest risk of Bipolar Disorder (Fig. 1b and Supplementary Table S5).

The results of comparable analysis for ASD, ADHD and Major Depressive Disorder are presented in Supplementary Table S6. No significant interaction effects were observed for ASD, ADHD or Major Depressive Disorder, consistent with the linear regression analysis. In addition, a sensitivity analysis for the Schizophrenia ANOVA result is presented in the Supplementary Information exploring the effect of varying the gestational age cut-off (Supplementary Methods and Results and Supplementary Tables S7a,b).

Finally, our analysis confirmed that cognitive outcome at four shows a positive association with gestational age in our cohort ($\beta = 0.039$, $se = 0.014$, $p = 5.5 \times 10^{-3}$) including sex and socio-economic status as covariates. None of the psychiatric genetic risk scores showed a significant association with cognition at age four (Supplementary Methods and Results and Supplementary Table S10), consistent with previous work in the literature⁸.

Discussion

In this work we tested for a possible interaction effect between psychiatric polygenic risk for five different psychiatric disorders and gestational age at birth on cognition at age four. We identified a significant interaction effect between genetic risk for both Bipolar Disorder and gestational age at birth and genetic risk for Schizophrenia and gestational age at birth. Those individuals with a greater genetic risk burden for either Bipolar Disorder or Schizophrenia appear to be more vulnerable to the adverse effects of birth at earlier gestational age on cognitive outcome. No significant interaction effect was observed for genetic risk for ASD, ADHD or Major Depressive Disorder and gestational age at birth.

The similar results obtained for both Bipolar Disorder and Schizophrenia are consistent with evidence for shared genetic aetiology for these disorders. In our cohort there is a significant correlation between genetic risk for Bipolar Disorder and Schizophrenia ($r = 0.266$, $p < 2.2 \times 10^{-16}$, Supplementary Table S8) which accords with results in the literature. In a large cohort, examining several disorders of the brain including psychiatric

pathologies, the most robust correlation was observed between genetic risk for Schizophrenia and Bipolar Disorder⁹ (correlation $r = 0.6808$, $p = 2.13 \times 10^{-230}$).

The observed correlation between genetic risk for Bipolar Disorder and Schizophrenia in our cohort motivated us to undertake an exploratory analysis looking at whether the interaction terms for Schizophrenia and Bipolar Disorder had an independent contribution to the regression model for cognition. We did this using stepwise linear regression and the final model included interaction terms for both pathologies (Supplementary Methods and Results and Supplementary Table S9).

Many studies in the literature have demonstrated an association between preterm birth and poorer cognitive outcome^{10–12}. This trend was replicated in our smaller dataset where we observed a positive association between cognition at four and gestational age at birth. Despite this consistent observation there remains a great deal of variability in the neurodevelopmental outcomes of preterm infants and much that is not understood about the underlying pathophysiology. The extent to which common genetic variation may play a role in understanding why some infants appear to be at greater risk from the environmental stress of prematurity remains largely unexplored. This work starts to address this question by asking if the effect of prematurity on cognitive outcome is dependent on genotype, and specifically psychiatric risk burden. In doing so we hope to be able to improve assessment of an individual's risk of cognitive impairment and better understand the pathophysiology underling the diverse neurodevelopmental outcomes observed in preterm infants. This will ultimately enable more effective identification of those infants at greatest risk of poorer outcomes and allow us to target therapeutic interventions appropriately.

The focus on genetic risk for psychiatric pathologies was motivated by previous work showing abnormal lenticular development in preterm infants with a greater genetic risk burden for psychiatric disease⁷. Furthermore, preterm birth is known to confer an increased risk of psychopathology. Preterm birth is associated with an increased risk of psychiatric hospitalization in adulthood across a range of psychiatric disorders³; the greatest increased risk is observed for Bipolar Disorder, followed by non-affective psychosis, which includes Schizophrenia and Schizoaffective Disorder.

Gene–environment interactions are the result of individuals responding differently to environmental stimuli, depending on their genotype. Many studies have explored gene–environment interactions for psychiatric pathologies¹³ and much of this work has focussed on understanding how stressful life events interact with genotype to affect risk of psychiatric conditions. The work of Capsi and colleagues¹⁴, which focused on the interaction between a functional candidate polymorphism in the serotonin transporter gene (5-HTT) and stressful life events, provides an early example of such a study. The authors showed that the level of depressive symptoms experienced by individuals in response to stressful life events was dependant on the allele they carried at this polymorphism.

Whilst early studies largely focussed on single polymorphisms, evidence suggests that a polygenic approach is likely to be more fruitful in understanding how genetic risk for psychiatric pathologies may increase susceptibility to environmental risk factors. More recent studies frequently investigate gene–environment interactions on a genome-wide scale, leveraging results from genome-wide association studies to construct polygenic scores. It is hoped that this approach could reveal interactions that may have gone undetected using candidate SNP analysis.

Gene–environment interaction studies exploring Schizophrenia have largely focussed on stressful life events¹⁵ and early life adversity^{16,17} with a few studies exploring perinatal factors. Studies with a perinatal focus include work indicating a significant interaction between maternal CMV infection and a polymorphism in the CTNNA3 gene on schizophrenia risk¹⁸ and work indicating that low birth weight predicts poorer academic and physical performance in individuals at high risk of Schizophrenia but not for those at low risk¹⁹.

Studies exploring Bipolar Disorder have followed similar themes to those for Schizophrenia focussing on environmental exposure to stressful life events¹³. The majority of work on Bipolar Disorder has been using candidate genes and to the best of our knowledge there have been no previous studies using summary statistics for Bipolar Disorder from the Psychiatric Genomics Consortium.

Limitations. Preterm birth itself may be a genetically inherited trait. However, it is difficult to know the extent to which this could contribute to confounding in this work. Despite evidence for a modest contribution of genetics to the risk of preterm birth, to-date, there exist few robust genetic associations identified through genome-wide association studies for prematurity. A large recent study identified two intergenic loci associated with preterm birth but failed to replicate these results²⁰.

Preterm birth is complex with heterogenous aetiology. Data that would allow us to distinguish between different aetiologies such as information on maternal infection or maternal pathology were not available in this study. It is possible that genetic risk alleles that confer an increased risk of Schizophrenia or Bipolar Disorder might also confer an increased risk of pathologies associated with preterm birth (e.g. pre-eclampsia, cervical insufficiency). Whilst we believe that such shared genetic aetiology, if present, is unlikely to fully explain our results, we cannot exclude the possibility that it could introduce confounding.

We have explored the possibility that genetic risk for psychiatric disease is associated with gestational age at birth within our cohort. For the five psychiatric pathologies examined in this study we found no significant association between genetic risk for psychiatric disease and gestational age at birth (Supplementary Methods and Results and Supplementary Table S11). Our result corroborates results in the literature: no association was found between genetic risk for ASD, ADHD or Schizophrenia and prematurity in a cohort of 7921 mothers and children from the Avon Longitudinal Study of Parents and Children²¹. That said, we cannot exclude the possibility that we are underpowered to detect such an association.

Environmental factors might additionally convey their influence through epigenetic mechanisms. There is growing evidence to suggest a role of DNA methylation in mediating the effects of preterm birth on future outcomes. For example, recent work in a cohort of 3648 newborns identified numerous differentially methylated

CpG sites in relation to gestational age at birth²². Further exploration of epigenetic modification associated with prematurity would be a relevant future study and one we hope to undertake using the Developing Human Connectome Cohort.

Finally, this study focused on cognitive outcome at age four. We hypothesised that the interaction between genetic risk for psychiatric disease and gestational age at birth would explain a significant fraction of the variability in cognitive outcome at age four. Work looking at preterm outcomes has shown stability of impairment in early childhood and confirmed that the most commonly observed impairment is in developmental or cognitive function². In future work we hope to look at whether or not our results extend to cognitive outcome at school age (7 and 12). As children grow older, smaller differences in ability might be easier to measure but they will similarly be subjected to a larger diversity of additional environmental exposures.

Summary. In summary, individuals with greater genetic risk for either Schizophrenia or Bipolar Disorder were shown to be more vulnerable to the negative effects of birth at earlier gestational age on cognitive outcome. The pathophysiology underlying preterm neurodevelopmental impairment may therefore be vulnerable to genetic variability associated with adult psychiatric disease.

Materials and methods

Subjects. Participants were drawn from the Twins Early Development Study (TEDS). Between 1994 and 1996 TEDS recruited over 15,000 twin pairs born in England and Wales, who have been assessed in multiple waves across their development up until the present date. Roughly 10,000 twin pairs still actively contribute to TEDS, providing genetic, cognitive, psychological and behavioural data. TEDS participants and their families are representative of families in the UK²³. Written informed consent was obtained from parents prior to data collection. Project approval was granted by King's College London's Research Ethics Committee for the Institute of Psychiatry, Psychology and Neuroscience (PNM/09/10-104) and the research was performed in compliance with the Declaration of Helsinki.

This study included 4934 unrelated individuals of European ancestry from the TEDS cohort. The dataset was subject to standard exclusions. This included exclusion of individuals with serious medical conditions which would likely affect their ability to take part in the TEDS assessments or conditions known to be associated with significant cognitive impairment. It also included exclusion of individuals in whom extreme circumstances were identified in the perinatal period such as prolonged hospital admission following birth. An initial filtering was then undertaken to select a cohort of unrelated individuals; this involved selection of all the unpaired twins plus one twin from each genotyped pair. Individuals were retained that had genotyping for the GPS analysis, data for cognitive outcome at age four, recorded gestational age at birth and recorded parental socio-economic status.

The final sample comprised 4934 individuals. Of these, 1858 were from monozygotic (MZ) twin pairs and 3076 were from dizygotic (DZ) twin pairs (1596 same-sex DZ pairs and 1480 opposite-sex DZ pairs). 2066 of the 4934 individuals were born at less than 37 weeks completed gestation and 918 were born at or below 34 weeks gestation.

Genotyping. Genotyping was undertaken at two separate timepoints with two different genotyping platforms. 1744 of the 4934 individuals were genotyped on the Affymetrix GeneChip 6.0 SNP array at Affymetrix, Santa Clara (California, USA) using buccal cell DNA samples. Genotypes were generated at the Wellcome Trust Sanger Institute (Hinxton, UK) as part of the Wellcome Trust Case Control Consortium 2 (<https://www.wtccc.org.uk/ccc2/>). The remaining 3190 individuals were genotyped on HumanOmniExpressExome-8v1.2 arrays at the Molecular Genetics Laboratories of the Medical Research Council Social, Genetic Developmental Psychiatry Centre, using DNA that was extracted from saliva samples.

After quality control, 635,269 SNPs remained for Affymetrix GeneChip 6.0 genotypes, and 559,772 SNPs for HumanOmniExpressExome genotypes. Genotypes from the two platforms were separately phased using EAGLE²⁴, and imputed into the Haplotype Reference Consortium (release 1.1) through the Sanger Imputation Service before merging genotype data from both platforms (for more details see Selzern et al.²⁵).

The final data contained 7,363,646 genotyped or well imputed SNPs. Principal component analysis was performed on a subset of 39,353 common (MAF > 5%), perfectly-imputed (info = 1) autosomal SNPs, after stringent pruning to remove markers in linkage disequilibrium ($r^2 > 0.1$) and excluding high linkage disequilibrium genomic regions.

Phenotypic data. *Cognition aged four.* The standardised cognitive ability measure at four was a composite of 'Parent Report of Children's Abilities' (PARCA) including both a parent-reported and parent-administered element and an assessment of verbal and non-verbal ability. The parent administered assessment included an odd-one-out task, design drawing task and puzzle task and the parent-reported assessment examined conceptual knowledge. Non-verbal ability was assessed with a 48-item vocabulary test and verbal ability with a grammar assessment which looked at length of sentences and correct use of both '-est' words and 'but'. The parent administered assessments have been validated in a subsample²⁶. The cognitive ability measure was computed as the mean of the four standardised elements.

Family socio-economic status (SES). Socio-economic status is a composite of maternal age at birth of eldest child, the mean score of maternal and paternal highest education level, as well as the respondent's (mother or father) occupation, administered by the Standard Occupational Classification 2000 (Office for National Statistics, 2000) when the child was two, which was the first age of contact.

Genome-wide polygenic scores. We calculated genome-wide polygenic scores, which are the SNP effect-size weighted sums of the number of trait-associated alleles. The SNP weights were derived from summary statistics for the largest GWA studies available for five psychiatric disorders at time of computation (see Supplementary Table S1).

The genome-wide polygenic scores were calculated using the software LDpred²⁷ which re-weights the SNP effect sizes based on a prior on the effect size and the LD in the sample. To ease computational demands, we selected variants with an info score of one, resulting in 515,000 SNPs used for analysis. In this work, we applied a prior on the fraction of causal markers of one for all analyses, based on the assumption that all genetic markers contribute to trait development. All polygenic scores were statistically adjusted for the first five ancestry principal components, chip and plate using the regression method and were z-standardized (mean = 0, SD = 1).

Statistical analysis. Linear regression modelling was used to estimate the effect of the gene-environment interaction ($GPS_{psych} \times GA$) on cognitive outcome at age four, for five different psychiatric pathologies (ADHD, ASD, Bipolar Disorder, Major Depressive Disorder and Schizophrenia) separately (Eq. (1)). Prior to inclusion in the model, the genome-wide polygenic scores were adjusted for ancestry, genotype chip and plate. Each model included the covariates sex and socio-economic status. Evidence suggests socio-economic status is an important predictor of cognitive outcome in both children and adults and that preterm infants may be particularly vulnerable to the effects²⁸. Interaction terms for psychiatric genome-wide polygenic score (GPS_{psych}) with sex and socio-economic status as well as interaction terms for gestational age at birth (GA) with sex and socio-economic status were included to properly control for possible confounding effects²⁹.

$$\begin{aligned} Cognition_{age4} = & \beta_0 + \beta_1 GPS_{psych} + \beta_2 GA + \beta_3 (GPS_{psych} \times GA) + \beta_4 (sex) \\ & + \beta_5 (SES) + \beta_6 (GPS_{psych} \times sex) + \beta_7 (GPS_{psych} \times SES) \\ & + \beta_8 (GA \times sex) + \beta_9 (GA \times SES) + \varepsilon \end{aligned} \quad (1)$$

The aim of this analysis was to investigate the effect of the interaction term $GPS_{psych} \times GA$ on cognitive outcome. We therefore explored whether the estimate of the coefficient β_3 of the interaction term deviated significantly from zero. P-values lower than the significance level $\alpha = 0.05/5 = 0.01$ were considered significant to account for the family-wise error rate using the Bonferroni method, correcting for the five psychiatric pathologies explored. Calculations were implemented with “lm” function in R (<https://www.r-project.org/>). Standard errors for the beta regression coefficients were estimated using the bootstrap technique with 10,000 bootstrap samples with random sampling of individuals with replacement.

As a further exploration of our data, we looked at the extremes of the data undertaking a two-way analysis of variance (ANOVA) comparing mean cognition between the extreme quintiles of polygenic risk for individuals born at or below 34 weeks gestation and term-born ($> = 37$ weeks) individuals for the psychiatric pathologies that yielded significant results in our initial analysis. Similar analysis for the non-significant pathologies is presented in the Supplementary Information (Supplementary Table S6).

In addition to our interaction analysis we sought to confirm that gestational age in our cohort predicted cognitive outcome at four. This was investigated using linear regression modelling, including both sex and socio-economic status as covariates (Eq. (2)).

$$Cognition_{age4} = \beta_0 + \beta_1 GA + \beta_2 (sex) + \beta_3 (SES) + \varepsilon \quad (2)$$

For completeness we also explored a possible association between cognition at four and genetic risk for each of the five psychiatric pathologies examined, details can be found in the Supplementary Analysis (Supplementary Methods and Results and Supplementary Table S10).

Ethics. Written informed consent was obtained from parents prior to data collection. Project approval was granted by King’s College London’s Research Ethics Committee for the Institute of Psychiatry, Psychology and Neuroscience (PNM/09/10-104). The research was performed in compliance with the Declaration of Helsinki.

Data availability

The Twins Early Development Study dataset is a controlled-access dataset that is accessible through collaboration with core TEDS researchers. A full explanation of the data access policy can be found at <http://www.teds.ac.uk/research/collaborators-and-data/teds-data-access-policy>.

Received: 15 December 2020; Accepted: 29 April 2021

Published online: 01 June 2021

References

1. Blencowe, H. *et al.* National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. *Lancet* **379**, 2162–2172 (2012).
2. Moore, T. *et al.* Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: The EPICure studies. *BMJ* **345**, 7961 (2012).
3. Mackay, D. F., Smith, G. C. S., Dobbie, R. & Pell, J. P. Gestational age at delivery and special educational need: Retrospective cohort study of 407,503 schoolchildren. *PLoS Med.* **7**, 1–10 (2010).
4. Johnson, S. & Marlow, N. Preterm birth and childhood psychiatric disorders. *Pediatr. Res.* **69**, 22–28 (2011).
5. Nosarti, C. *et al.* Preterm birth and psychiatric disorders in young adult life. *Arch. Gen. Psychiatry* **69**, 1–8 (2012).

6. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *J. Child Psychol. Psychiatry Allied Discip.* **55**, 1068–1087 (2014).
7. Cullen, H. *et al.* Polygenic risk for neuropsychiatric disease and vulnerability to abnormal deep grey matter development. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-38957-1> (2019).
8. Krapohl, E. *et al.* Phenome-wide analysis of genome-wide polygenic scores. *Mol. Psychiatry* <https://doi.org/10.1038/mp.2015.126> (2015).
9. Anttila, V. *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
10. Brydges, C. R. *et al.* Cognitive outcomes in children and adolescents born very preterm: a meta-analysis. *Dev. Med. Child Neurol.* **60**, 452–468 (2018).
11. Twilhaar, E. S., De Kieviet, J. F., Aarnoudse-Moens, C. S. H., Van Elburg, R. M. & Oosterlaan, J. Academic performance of children born preterm: A meta-analysis and meta-regression. *Arch. Dis. Child. Fetal Neonatal Ed.* **103**, 322–330 (2018).
12. Pettinger, K. J. *et al.* Starting school: Educational development as a function of age of entry and prematurity. *Arch. Dis. Child.* **105**, 160–165 (2020).
13. Musci, R. J., Augustinavicius, J. L. & Volk, H. Gene-environment interactions in psychiatry: Recent evidence and clinical implications. *Curr. Psychiatry Rep.* **21**, 1–10 (2019).
14. Caspi, A. *et al.* Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389 (2003).
15. Mandelli, L., Toscano, E., Porcelli, S., Fabbri, C. & Serretti, A. Age of onset in schizophrenia spectrum disorders: Complex interactions between genetic and environmental factors. *Psychiatry Investig.* **13**, 247–249 (2016).
16. Hoffmann, C. *et al.* Exploring the moderating effects of dopaminergic polymorphisms and childhood adversity on brain morphology in schizophrenia-spectrum disorders. *Psychiatry Res.* **281**, 61–68 (2018).
17. Trotta, A. *et al.* Interplay between schizophrenia polygenic risk score and childhood adversity in first-presentation psychotic disorder: A pilot study. *PLoS ONE* **11**, 1–14 (2016).
18. Børglum, A. D. *et al.* Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci. *Mol. Psychiatry* **19**, 325–333 (2014).
19. Forsyth, J. K. *et al.* Genetic risk for schizophrenia, obstetric complications, and adolescent school outcome: Evidence for gene-environment interaction. *Schizophr. Bull.* **39**, 1067–1076 (2013).
20. Rappoport, N. *et al.* A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth. *Sci. Rep.* **8**, 1–11 (2018).
21. Leppert, B. *et al.* Association of maternal neurodevelopmental risk alleles with early-life exposures. *JAMA Psychiat.* **76**, 834–842 (2019).
22. Merid, S. K. *et al.* Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age. *Genome Med.* **12**, 1–17 (2020).
23. Rimfeld, K. *et al.* Twins early development study: A genetically sensitive investigation into behavioral and cognitive development from infancy to emerging adulthood. *Twin Res. Hum. Genet.* **22**, 508–513 (2019).
24. Loh, P. R. *et al.* Reference-based phasing using the Haplotype reference consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
25. Selzam, S. *et al.* Evidence for gene-environment correlation in child feeding: Links between common genetic variation for BMI in children and parental feeding practices. *PLoS Genet.* **14**, 1–19 (2018).
26. Oliver, B. *et al.* The validity of a parent-based assessment of cognitive abilities in three-year olds. *Early Child. Dev. Care* **172**, 337–348 (2002).
27. Vilhjalmsón, B. J. *et al.* Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
28. Wolke, D. Is social inequality in cognitive outcomes increased by preterm birth-related complications?. *JAMA Netw. Open* **2**, 192902 (2019).
29. Keller, M. C. Gene × environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biol. Psychiatry* **75**, 18–24 (2014).

Acknowledgements

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant to Robert Plomin from the UK Medical Research Council (MR/M021475/1 and previously G0901245). The research leading to these results has also received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ grant agreement n° 602768 and ERC grant agreement n° 295366. Harriet Cullen is funded by a National Institute for Health Research (NIHR), Academic Clinical Lectureship at King's College, London for this research project. The work was also supported by the NIHR Biomedical Research Centre at Guy's and St Thomas' NHS Trust, the Wellcome/EPSCRC Centre for Medical Engineering at King's College London and the MRC Centre for Neurodevelopmental Disorders, Kings College London. This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Author contributions

Study conceptualisation and design: H.C., S.S., R.P., A.D.E.; data analysis: H.C., S.S., K.D.; writing and figure prep: H.C., S.S.; reviewing manuscript: H.C., S.S., K.D., R.P., A.D.E.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90045-5>.

Correspondence and requests for materials should be addressed to H.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021