The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry 63:10 (2022), pp 1140–1152 doi:10.1111/jcpp.13657

Polygenic scores for schizophrenia and major depression are associated with psychosocial risk factors in children: evidence of gene–environment correlation

Sandra Machlitt-Northen,[1] Robert Keers,[1] Patricia B. Munroe,[2] David M. Howard,[3,4] Vassily Trubetskoy,[5] and Michael Pluess[1]

[1]Department of Biological and Experimental Psychology, Queen Mary University of London, London, UK; [2]Department of Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK; [3]Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK; [4]Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK; [5]Department of Psychiatry and Psychotherapy, Universitätsmedizin Berlin Campus Charité Mitte, Berlin, Germany

Background: Whilst genetic and environmental risk factors for schizophrenia (SCZ) and major depressive disorder (MDD) have been established, it is unclear whether exposure to environmental risk factors is genetically confounded by passive, evocative or active gene–environment correlation (rGE). **Study Objective:** This study aims to investigate: (a) whether the genetic risk for SCZ/MDD in children is correlated with established environmental and psychosocial risk factors in two British community samples, the 1958 National Child Development Study (NCDS) and the Millennium Cohort Study (MCS), (b) whether these associations vary between both psychopathologies, and (c) whether findings differ across the two cohorts which were born 42 years apart. **Methods:** Polygenic risk scores (PRS) from existing large genome-wide associations studies (GWAS) were applied to test the correlation between the child genetic risk for SCZ/MDD and known environmental risk factors. In addition, parental and child genetic data from MCS were used to distinguish between passive and evocative rGE. **Results:** The child polygenic risk for SCZ and MDD was correlated with single parenthood in MCS. Moreover, the lack of father's involvement in child care was associated with the genetic risk for SCZ in NCDS. However, we also found associations between several indicators of low socioeconomic status and heightened genetic risk for MDD in children in both cohorts. Further, the genetic risk for MDD was associated with parental lack of interest in the child's education in NCDS as well as more maternal smoking and less maternal alcohol consumption during childhood in MCS. According to sensitivity analyses in MCS (controlling for parental genotype), more than half of our significant correlations reflected passive rGE. **Conclusions:** Findings suggest that several established environmental and psychosocial risk factors for SCZ and MDD are at least partially associated with children's genetic risk for these psychiatric disorders. **Keywords:** Environment; schizophrenia; major depressive disorder; genetics.

Introduction

According to classic twin and family studies, schizophrenia (SCZ) and major depression (MDD) are associated with both genetic and environmental risk factors which are likely intertwined in complex ways (Kendler & Karkowski-Shuman, 1997; Sullivan, Kendler, & Neale, 2003). In this article, we aim to investigate this gene–environment interplay by testing: (a) whether established environmental risk factors for SCZ/MDD in childhood are correlated with the genetic risk of these disorders, measured with polygenic risk scores (PRS) derived from published large genome-wide association studies (GWAS), (b) whether patterns of gene–environment correlations differ between SCZ and MDD, as well as (c) between two British cohorts that are 42 years apart.

Environmental risk

Various environmental and psychosocial factors in early childhood are known to be associated with SCZ and MDD. For instance, prenatal and postnatal risk factors for SCZ include maternal smoking during pregnancy (Sacker, Done, Crow, & Golding, 1995) as well as obstetric complications such as prenatal infection (Brown & Derkits, 2010) and foetal hypoxia (Cannon et al., 2000). Moreover, low birth weight (Abel et al., 2010; Alati et al., 2007) and short gestational age (Chiu et al., 2019; Jones, Rantakallio, Hartikainen, Isohanni, & Sipila, 1998) emerged as established risk factors for both psychopathologies. Importantly, it appears that markers of material disadvantage in early childhood are of particular relevance for the risk of SCZ and MDD (Burns, Tomita, & Kapadia, 2014; Nasir & Bloch, 2021), given that low socioeconomic status (SES) at birth (Freeman et al., 2016; Harrison, Gunnell, Glazebrook, Page, & Kwiecinski, 2001) and low parental educational attainment (Cohen,

Conflict of interest statement: See Acknowledgements for full disclosures.

Nussbaum, Weintraub, Nichols, & Yen, 2020; Keefe, Eesley, & Poe, 2005) are known to be involved in the development of both disorders.

Genetic risk

Family and twin studies suggest that SCZ and MDD have a substantial heritability in adults of 79%–81% (Hilker et al., 2018; Sullivan et al., 2003) and 37%–39% (Kendler & Prescott, 1999; Sullivan, Neale, & Kendler, 2000), respectively. Whilst few studies have assessed the heritability of SCZ in child samples, which is further complicated by the prodromal phase in early adolescence (Kahn et al., 2015), twin studies suggest that the genetic influence of depressive symptoms in children is small but increases in adolescence (Rice, 2010). This is supported by a twin study of female adolescents aged 12–19 years from the Missouri Adolescent Female Twin Study (MOAFTS) where the heritability for MDD was estimated to be similar to that of adults at 40% (Glowinski, Madden, Bucholz, Lynskey, & Heath, 2003).

Furthermore, recent genome-wide association studies propose that the identified genetic component is made up of thousands of single nucleotide polymorphisms (SNPs) with small effect sizes, which can be aggregated into polygenic risk scores (PRS) (Dudbridge & Newcombe, 2015). Whilst SCZ and MDD are distinct psychiatric disorders, empirical studies suggest that both share some of their genetic architecture, with GWAS findings from the SCZ Workgroup and the MDD Consortium having identified 108 and 44 independent risk loci, respectively (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Wray et al., 2018). Moreover, both disorders share six overlapping loci, located close to the Transcription Factor 4 (TCF4) protein coding gene which is crucial for normal brain development, highlighting a substantial genetic correlation (r_g = .34) between them (Wray et al., 2018).

Furthermore, evidence from molecular genetics as well as twin and family studies stresses the genetic link between childhood and adulthood psychiatric disorders (Nivard et al., 2017). For instance, one recent PRS study in 9,912 adolescents from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort suggested that the genetic liability to SCZ often manifests itself as negative symptoms and anxiety during adolescence (Jones et al., 2016). On the other hand, Kwong et al. (2021) found that PRSs for MDD, depression, neuroticism and anxiety were correlated with increased depressive symptoms in adolescence as well as early adulthood in over 6,000 participants from ALSPAC, emphasising that genetic susceptibility may play a part in adolescence depression.

Gene–environment correlations

Considering genes and environment rarely work in isolation, one important type of interplay between both is gene–environment correlation (rGE) which describes how an individual's genetic disposition can influence the exposure to particular environments (Plomin, DeFries, & Loehlin, 1977). *Passive* rGE occurs before birth when the biological parents pass on genetic traits to their children whilst also providing a specific environmental context (Jaffee & Price, 2008). *Evocative* rGE occurs from birth onwards when a genetic predisposition gives rise to particular behaviour that evokes a specific response from the environment, whilst *active* rGE refers to an individual actively seeking out particular environments due to their genetic predisposition usually occurring later in life (Jaffee & Price, 2008).

Findings from twin and adoption studies, which infer the heritable contribution from familial correlations, provide substantial empirical evidence that environments are heritable (Jaffee & Price, 2007). Although molecular genetic analyses, which measure the genotype directly, still make up a smaller segment of rGE research, these studies further substantiate the existence of rGE (Jaffee & Price, 2007). One of these recent GWAS analyses that investigated the covariation between environmental exposures and children's genetic propensity for several outcomes found that the genetic risk for SCZ in children was associated with increased paternal age, reflecting passive rGE (Krapohl et al., 2017). Moreover, Ensink et al. (2020) reported possible passive rGEs in children from a Dutch birth cohort for a range of existing PRSs: the PRS for SCZ was negatively correlated with maternal education and the PRS for MDD positively correlated with mother's prenatal anxiety. However, whilst first well-powered and longitudinal rGE studies featuring PRSs are now emerging, few study designs are able to disentangle the cross-generational relationships of genes and the environment further. A clear understanding of the biological pathways of our genes to behaviours and thus associated environments has important consequences for the prevention or treatment of psychopathologies (Jaffee & Price, 2007, 2012). In order to distinguish between passive or evocative rGE, Krapohl et al. (2017) argue that parental as well as child genotypes and phenotypes need to be available in the same sample. Our study will address this gap by drawing on genetic data and environmental measures from both children and parents in order to differentiate between passive and evocative rGE.

Study objectives

The aims of this study are to test for the presence of rGEs in childhood and investigate the nature of rGE (i.e. passive vs. evocative) by controlling for parental genotype in one of the two cohorts. Bearing in mind that PRS studies utilise GWAS results obtained from adult cohorts, it is also important to clarify whether these adult-based PRS for SCZ and MDD apply to children in the general population.

Furthermore, we wanted to test whether rGE differs between SCZ and MDD and between the two community cohorts from different generations given that gene–environment interplay may change over time as a result of cultural shifts in risk and behaviour.

Hypotheses

Based on existing literature, we hypothesised that established environmental risk factors for SCZ or MDD would correlate with their genetic risk in children. However, given that genetic overlap is not complete between SCZ and MDD, we expected rGE results to differ between disorders. Furthermore, we expected that a subset of observed rGE in children will be accounted for by the parental genotypes through passive rGE. Finally, we anticipated that associations would differ across the two generational cohorts given societal changes in living conditions and health behaviours.

Methods

Participants

Participants were drawn from two cohort studies: the Millennium Cohort Study (MCS) and the 1958 National Child Development Study (NCDS). The MCS includes 18,827 children who were born in either England and Wales between September 2000 and August 2001 or in Scotland and Northern Ireland between November 2000 and January 2002 (Connelly & Platt, 2014). A total of 692 new families were added at age 3 bringing the total to 19,517 cohort members of which 9,894 (51.4%) were male and 15,638 (81.3%) white (self-reported) (Staatz, Kelly, Lacey, & Hardy, 2021). The first survey was completed by the parents with later surveys including parents, teachers and cohort member reports. Ethics approval [MREC/01/6/19, MREC/03/2/022, 05/MRE02/46, 07/MRE03/32, 11/YH/0203, 13/LO/1786] was received from the London Multicentre Research Ethics Committee (MREC) prior to each survey (Shepherd & Gilbert, 2019). Participating parents and teachers provided informed consent. Wave 6 included the collection of 23,336 saliva samples from cohort members aged 14 years of age and their biological parents for DNA extraction, resulting in 4,533 mother, child, father "trios" for which ethics approval was obtained from London-Central REC (13/LO/1786) (Fitzsimons et al., 2020).

The NCDS comprises 17,415 unrelated individuals (Power & Elliott, 2006). All individuals were born in England, Wales or Scotland in a single week in March 1958 (Bann, Johnson, Li, Kuh, & Hardy, 2018). Over 98% of initial participants were of white ethnic background, but the dataset was later augmented with immigrants born within the same reference week (Bann et al., 2018; Power & Elliott, 2006). The initial birth survey was completed by the midwife and includes data from clinical records, whereas data at ages 7, 11 and 16 years were collected from parents and teachers. A bio-medical survey, which included blood samples for DNA extraction, was conducted between 2002 and 2004 on 9,293 individuals aged 44 to 46 years of age. All required ethical approvals were obtained [01/1/44; 08/H0718/29; 12/LO/2010], including for the biomedical survey (Centre for Longitudinal Studies, 2014). NCDS requested consent from parents or respondents at each wave, with participants required to provide additional written informed consent for the bio-medical assessments at age 44 (Centre for Longitudinal Studies, 2014).

For both cohorts, we had access to genome-wide as well as psychosocial data. For MCS, we used data from six waves (9 months in 2001 to 14 years in 2015), and for the NCDS, we included four waves (birth in 1958 to 16 years in 1974).

Measures

Environmental factors. The following available and established environmental and psychosocial risk factors for SCZ or MDD were included: low birth weight, parity, short gestational period, mother's and father's age at birth, as well as maternal smoking prior and during pregnancy, parental substance abuse, such as smoking and alcohol consumption, socioeconomic status (SES) and indicators thereof, such as unemployment, financial difficulties, housing issues, tenure of the house (owned or rented), number of bedrooms in the family home, and whether the child received free school meals. Moreover, we also selected maternal and paternal interest in the child's education, paternal involvement in child care, whether the mother or father is taking the child for walks or to the park or is reading to the child, given that educational attainment is an established environmental risk factor for both psychopathologies. Finally, we also considered parental marital status, including divorce or separation and domestic tension (see Appendix S1 for more details).

Genetic data. For MCS, we used genome-wide data from 21,324 individuals (8,201 children and 13,123 parents) genotyped on Illumina's Infinium global screening array (GSA)-24 v1.0 (Fitzsimons et al., 2020). For NCDS, we used SNP data from three arrays: 1,502 individuals genotyped on the Affymetix 500 k 1.2 M for The Wellcome Trust Case Control Consortium 1 (WTCCC1) (Wellcome Trust Case Control Consortium, 2007), 2,922 individuals genotyped using the Illumina 1.2 M array for Wellcome Trust Case Control Consortium 2 (WTCCC2) (Wellcome Trust Case Control Consortium 2, n.d.) and 2,592 individuals genotyped on Infinium Humanhap 550 k v3 chips for the Type 1 Diabetes Genetics Consortium (T1DGC) (Barrett et al., 2009).

Quality control and imputation. Quality control (QC) was carried out separately for MCS (21,446 samples) and NCDS (2,922; 2,592 and 1,502 samples from WTCCC2, T1DGC and WTCCC1, respectively).

First, for MCS only, the dataset was split into children and parents before creating a linkage-disequilibrium (LD) pruned set of markers for each subset in order to calculate genome-wide identical-by-state (IBS) sharing using PLINK 1.9 (Chang et al., 2015; Purcell et al., 2007). Individuals were then clustered into homogeneous groups (k = 14) through a multi-dimensional scaling analysis. The homogenous clusters closest to European ancestry, when overlaid with reference individuals from the 1,000 Genomes Project (Auton et al., 2015), were used as the European ancestry subsets (7,025 children and 11,269 parents) with all other clusters combined into non-European subsets (1,176 children and 1,852 parents).

Using PLINK 1.9, QC for the MCS and NCDS subsets were performed according to Coleman et al. (2016) which included the removal of duplicates, minor allele frequencies (MAF) of <1%, variants or samples with missing data or data of low quality were removed in an iterative manner until a final 99% threshold was reached (80–99% threshold in 1% intervals for MCS and 90–99% threshold in 1% intervals for NCDS), and SNPs with a Hardy–Weinberg equilibrium p-value of $<1 \times 10^{-5}$. SNPs were pruned for linkage disequilibrium (LD) ($r^2 < 0.2$), high-LD or nonautosomal regions, and individuals with mismatching phenotype and genetic sex were excluded. Identical-by-descent checks (pi-hat < 0.1875), population stratification and ancestry groupings were run using

EIGENSTRAT (Price et al., 2006) and PERL (Patterson, Price, & Reich, 2006) for the top 100 principal components. Outliers were compared against the 1,000 Genomes Project reference panel (Auton et al., 2015) before being removed. Further checks for unusual genome-wide heterogeneity were also performed before flipping reverse strand SNPs and discarding ambiguous SNPs using SNPFLIP v0.0.6 (Bakken Stovner, 2017) and removing any SNPs with allele frequency mismatches (Auton et al., 2015). NCDS data were lifted over using liftOverPlink (Ritchie, 2014) from B35 to B37 for WTCCC1 and B36 to B37 for WTCCC2 and T1DGC. The genetic data were imputed using the 1,000 Phase 3 Genomes Project reference panel (Auton et al., 2015) before performing postimputation QC. Postimputation output filters for imputation quality ($R^2 > .8$) and posterior genotype probability imputation confidence (GP threshold of >.8) were applied using bcftools (Danecek et al., 2021) prior to converting the VCF files to PLINK format. Failed or duplicated SNPs, MAF (<5%), and missing SNPs or individuals (<99%) were excluded. Individual subsets were merged back together for each cohort and triallelic sites were removed. The final MCS and NCDS datasets included 6,634,361 and 6,398,736 variants in 18,476 (7,280 children, 4,322 Fathers, 6,874 Mothers) and 5,288 individuals, respectively. Principal component analysis was re-run on both LD-pruned cohorts using smart.pca before selecting the top principal components out of 100 which explained the majority of the variance in order to be included as covariates in the regression analysis. See Appendix S2 for more details.

Polygenic risk scores. For both cohorts, individual PRS were calculated in PRSice (Euesden, Lewis, & O'Reilly, 2015) at seven thresholds (0.01, 0.1, 0.2, 0.3, 0.4, 0.5 and 1) based on GWAS results from the SCZ Workgroup (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and the MDD Consortium (Wray et al., 2018). Given that NCDS was used as a control sample by both consortia, revised GWAS results were obtained which excluded NCDS data.

Data analysis

All data analysis was performed in Stata v12.1 (StataCorp., 2011). Descriptive statistics were run for both cohorts and each genotyped subsample for every environmental measure at each time point. We used unpaired *t*-tests or chi-square tests for polytomous/continuous variables and binary variables, respectively, to assess whether the genotyped subsample was representative of the whole cohort sample. To assess multicollinearity, correlation matrices with pairwise correlation coefficients were analysed for all indicators of SES (Appendices S3 and S4 for more details).

Logistic and linear regressions were run for variables which were only available at a single time point. Environmental variables that were measured repeatedly at different waves were combined into either logistic or linear mixed effects longitudinal models or random effects longitudinal models.

For MCS, all regressions were run using child PRSs (7,280 individuals). Year of data collection (MCS), current age (NCDS), sex and the top 8 and 5 principal components for NCDS and MCS, respectively, which account for population stratification, were used as covariates in all regression models for both cohorts. Results were corrected for multiple testing applying Bonferroni correction (corrected $p = \alpha/\text{number of environments}$: $5.81 \times 10^{-4} = 0.05/86$). Correlations between the child PRS and environmental risk factor were considered statistically significant if at least one of the seven thresholds was below the corrected Bonferroni *p*-value.

According to power analyses with G*Power v3.1 (Faul, Erdfelder, Buchner, & Lang, 2009), both samples were sufficiently powered for all variables except for father's interest in the child's education at wave 4 of MCS and father's interest in the child's education at ages 7, 11 and 16 in NCDS.

The following sensitivity analyses were performed for all significant correlations: In MCS, maternal and/or paternal PRSs were added as covariates to logistic/linear regression models or mixed effects longitudinal models to allow for the distinction between passive or evocative rGE. In NCDS, we wanted to account for the possibility that findings could be confounded by the presence of clinical cases. Therefore, for SCZ we re-ran the analyses by excluding individuals with a self-reported diagnoses of SCZ, psychosis or hallucinations at age 55, whereas for MDD we removed individuals who reported depression as adults aged 55.

Additionally, to assess whether the resulting regression coefficients are statistically different from each other, we conducted an interaction analysis between all independent variables and (a) the SCZ or MDD symptoms in NCDS, and (b) the maternal/paternal PRS in MCS between the original analyses and the sensitivity analyses. We then used the resulting Wald Chi-squared test statistics of the PRS interaction terms to assess whether the coefficients are statistically different from each other. Moreover, we repeated the interaction analysis for any findings which matched between the two disorders by interacting the child PRS for SCZ and the child PRS for MDD with all independent variables in the regression models.

Results

Descriptive statistics

In MCS, several significant differences emerged between the included genetic subsample and the original cohort (for whom phenotypic data were available). The genetic subsample has a higher percentage of individuals who rented in wave 6, higher odds for mothers who smoked at wave 1, a greater proportion of individuals who fell into the SES class 1 category at waves 1, 3, 4 and a decreased proportion of individuals who were married at all waves. In the NCDS, individuals in our selected genetic sample had higher odds for renting at wave 7 compared to the original sample (for whom phenotypic data was available). More details are provided in Appendix S4.

rGE results for MCS

We identified only one significant rGE result between the genetic risk for SCZ in children and parental marital status in MCS, with parents of children with higher genetic risk for SCZ being more likely to be divorced, separated or widowed. However, we found several significant associations between the genetic risk for MDD and indicators of material disadvantage, including lower SES and living in rented accommodation. In addition, we obtained significant rGE results for the child's genetic risk for MDD and more maternal smoking and less alcohol intake, which both depict consumption behaviours of the mother at various timepoints during childhood, as well as parents being more likely to be divorced, separated or widowed (see Table 1 for all MCS results for PRS at three selected thresholds: 0.01,

0.5 and 1). According to our sensitivity analyses, the association between the child PRS and parental marital status for both psychopathologies as well as tenure and maternal smoking for MDD was confounded by parental genotype, whereas maternal alcohol consumption and SES in MDD were not (for more details see Appendix S5).

rGE results for NCDS

We identified an association between the lack of father's involvement in child care and the genetic risk for SCZ in NCDS. Similarly, to results in MCS, we identified several significant correlations between the genetic risk for MDD and lower SES, living in rented accommodation and lower number of bedrooms. Moreover, the genetic risk for MDD in children was correlated with heightened maternal and paternal lack of interest in the child's education (see Table 2 for all NCDS results for PRS at three thresholds: 0.01, 0.5 and 1). Sensitivity analyses suggested that our findings cannot be explained by the presence of clinical cases (for more details see Appendix S5).

Discussion

This study aimed to test whether the genetic risk for SCZ and MDD in children, measured with PRS, was correlated with established environmental and psychosocial risk factors. We also investigated whether rGE differed between the two disorders and whether detected rGEs are comparable across generations.

rGE for SCZ

We identified a statistically significant association for SCZ that involved a higher risk for parents being divorced, separated, or widowed in MCS. Not surprisingly, this correlation was partially confounded by the parental genotype and therefore most likely reflects passive rGE, whereby parents provide the family environment whilst also passing down their genes to their offspring. Moreover, we found an association between the father's lack of involvement in child care and the genetic risk for SCZ in NCDS. However, we were unable to disentangle which form of rGE was present in NCDS due to the lack of parental genotypes. Nevertheless, given that SCZ is highly heritable, it is surprising that we did not detect more rGEs. One reason may be the low base rate of SCZ in the general population (Jaffee & Price, 2007), although PRS should overcome this limitation at least partially.

rGE for MDD

Our study provides consistent evidence for significant associations between the genetic risk for MDD in children and various risk factors. Several points deserve further discussion: First, detected rGEs involve multiple markers of low SES in both cohorts. However, according to sensitivity analyses in MCS, only tenure was confounded by the parental genetic risk, most likely through passive rGE, whereas SES was not and could therefore be a consequence of the limiting symptoms of the disease itself. Second, many of the correlations involved parental behaviours (e.g. maternal smoking). Based on our sensitivity analyses, maternal smoking is associated with the child genetic risk for MDD through passive rGE, whereas maternal alcohol consumption appears to be at least partially explained by evocative rGE. Third, we also identified associations between the child PRS and psychosocial risk factors (e.g. maternal and paternal lack of interest in the child's education). Whilst we cannot disentangle the form of rGE in NCDS, it is possible that the parenting behaviour themselves is partly heritable through passive rGE. This is, for instance, confirmed by a recent study which suggested that the PRS for educational attainment was mediated by parents' cognitive abilities and self-control skills (Wertz et al., 2019).

Differences between SCZ and MDD

Whilst we observed eight significant rGEs between the PRS for MDD and environmental risk factors, only two rGE emerged for the genetic risk for SCZ. Although parental marital status was associated with the child genetic propensity for both psychopathologies in MCS, our comparison of regression coefficients indicated that the strength of the association was similar for both SCZ and MDD, possibly due to the partially shared genetic overlap. Results suggest that rGEs differ between disorders, with the genetic risk for MDD being more strongly associated with childhood environmental and psychosocial risk factors compared to SCZ. Interestingly, whilst SCZ is less prevalent and having a higher heritability than MDD, rGE in childhood appears less relevant for SCZ. Given the evidence from twin studies, it is also possible that rGE may play a stronger role in MDD due to the lower heritability. Any passive or evocative rGE in childhood may be influenced by the greater non-shared environments as opposed to the genetic portion of variance, making rGE more prominent in MDD.

Differences between cohorts

The third aim of this study was to investigate whether rGE findings differ across generations. Our findings indicate only a partial overlap of results between the two cohorts.

The genetic risk for SCZ and MDD was associated with single parenthood in MCS but not in NCDS whose participants were 40 years older. This

Table 1 Regression results of rGE in the MCS sample

	SCZ				MDD		
Environment	PRS threshold *z*-scored	Beta	95%CI	*p*-Value	Beta	95%CI	*p*-Value
SES	0.01	−0.01	−0.06 to 0.03	5.32E-01	−0.07	−0.10 to −0.03	**1.04E-04****
(5 Professional/managerial, 4 Intermediate, 3 Small employer/self-employed, 2 Lower supervisory/technical, 1 Semi-routine/routine)	0.5	−0.02	−0.09 to 0.05	5.44E-01	−0.1	−0.19 to −0.01	2.54E-02*
	1	−0.02	−0.09 to 0.05	5.79E-01	−0.09	−0.17 to −0.00	4.34E-02*
Finance issues	0.01	0.01	−0.14 to 0.16	8.59E-01	0.19	0.08–0.30	8.35E-04*
(0 = no, 1 = yes)	0.5	0.11	−0.12 to 0.34	3.43E-01	0.29	0.01–0.56	3.88E-02*
	1	0.09	−0.14 to 0.32	4.36E-01	0.21	−0.05 to 0.46	1.10E-01
Number of rooms	0.01	0	−0.04 to 0.04	9.51E-01	−0.05	−0.08 to −0.02	2.85E-03*
(continuous)	0.5	0	−0.06 – 0.07	9.29E-01	-0.07	−0.15 to 0.01	7.61E-02
	1	0	−0.06 to 0.07	8.80E-01	−0.05	−0.13 to 0.03	1.88E-01
Tenure	0.01	0.22	−0.04 to 0.47	9.91E-02	0.43	0.24–0.63	**1.14E-05****
(0 = owns,1 = rents)	0.5	0.28	−0.11 to 0.68	1.55E-01	0.62	0.14–1.10	1.20E-02*
	1	0.27	−0.13 to 0.66	1.87E-01	0.51	0.06–0.96	2.76E-02*
Mother's interest in child's education	0.01	0.16	−0.10 to 0.41	2.41E-01	0.1	−0.10 to 0.30	3.16E-01
(0 = interested, 1 = uninterested)	0.5	0.08	−0.31 to 0.48	6.81E-01	0.09	−0.43 to 0.61	7.33E-01
	1	0.07	−0.33 to 0.47	7.28E-01	0.09	−0.40 to 0.58	7.10E-01
Father's involvement in upbringing	0.01	0.26	0.01–0.50	4.05E-02	0.26	0.07–0.46	8.15E-03*
(0 = yes, 1 = no)	0.5	0.3	−0.08 to 0.67	1.22E-01	0.36	−0.20 to 0.92	2.03E-01
	1	0.29	−0.09 to 0.67	1.29E-01	0.32	−0.23 to 0.86	2.53E-01
Father's interest in child's education	0.01	0.12	−0.15 to 0.38	3.99E-01	0.12	−0.08 to 0.32	2.54E-01
(0 = interested, 1 = uninterested)	0.5	0.08	−0.33 to 0.48	7.09E-01	0.39	−0.17 to 0.94	1.73E-01
	1	0.09	−0.32 to 0.50	6.67E-01	0.41	−0.12 to 0.95	1.31E-01
Mother walks	0.01	−0.03	−0.20 – 0.14	7.51E-01	-0.1	−0.22 to 0.03	1.32E-01
(0 = weekly, 1 = monthly or less)	0.5	−0.17	−0.43 to 0.09	2.00E-01	−0.1	−0.42 to 0.23	5.61E-01
	1	−0.18	−0.44 to 0.09	1.90E-01	−0.15	−0.46 to 0.15	3.23E-01
Father walks	0.01	−0.06	−0.25 – 0.13	5.42E-01	-0.1	−0.25 to 0.05	1.82E-01
(0 = weekly, 1 = monthly or less)	0.5	−0.13	−0.43 to 0.16	3.85E-01	−0.21	−0.63 to 0.22	3.41E-01
	1	−0.15	−0.45 to 0.15	3.39E-01	−0.17	−0.57 to 0.23	4.10E-01
Smoking mother	0.01	0.08	0.02–0.15	1.67E-02*	0.2	0.15–0.26	**0.00E+00****
(0 = no, 1 = yes)	0.5	0.11	0.01–0.22	3.86E-02*	0.45	0.30–0.60	**7.10E-09****
	1	0.11	0.00–0.22	4.77E-02*	0.39	0.24–0.54	**1.61E-07****
Smoking father	0.01	0.19	−0.16 to 0.54	2.92E-01	0.17	−0.09 to 0.43	1.94E-01
(0 = no, 1 = yes)	0.5	0.12	−0.41 to 0.66	6.57E-01	0.37	−0.38 to 1.12	3.37E-01
	1	0.12	−0.42 to 0.66	6.64E-01	0.27	−0.44 to 0.98	4.64E-01
Gestational period	0.01	0.02	−0.03 to 0.06	4.82E-01	0.01	−0.02 – 0.04	5.60E-01
(continuous)	0.5	0.03	-0.04 to 0.09	3.97E-01	−0.07	−0.15 to 0.02	1.16E-01
	1	0.03	−0.04 to 0.10	4.13E-01	−0.04	−0.12 – 0.03	2.65E-01
Birth weight	0.01	0.02	−0.03 to 0.08	3.90E-01	0.01	-0.03 to 0.05	7.80E-01
(continuous)	0.5	0.08	0.00–0.16	4.22E-02*	−0.07	−0.18 to 0.04	2.05E-01
	1	0.08	0.00–0.16	4.91E-02*	−0.05	−0.16 to 0.05	2.95E-01
Marital status	0.01	0.1	0.06–0.15	**1.07E-05****	0.09	0.06–0.13	**2.17E-07****
(0 = married/civil partner, 1 = divorced/separated)	0.5	0.18	0.11–0.25	**8.36E-07****	0.2	0.10–0.29	**3.61E-05****
	1	0.17	0.10–0.24	**2.25E-06****	0.18	0.09–0.27	**4.96E-05****

(continued)

Table 1 (continued)

	SCZ				MDD		
Environment	PRS threshold z-scored	Beta	95%CI	*p*-Value	Beta	95%CI	*p*-Value
Alcohol mother	0.01	−0.04	−0.09 to 0.01	8.00E-02	−0.09	−0.13–−0.05	**3.08E-06****
(0 = less than 1–2 weekly, 1 = more than 1–2 weekly)	0.5	0.02	-0.05 to 0.09	6.06E-01	−0.26	−0.36 to −0.15	**2.91E-06****
	1	0.02	−0.06 to 0.09	6.81E-01	−0.25	−0.35 to −0.14	**3.67E-06****
Alcohol father	0.01	−0.12	−0.38 to 0.13	3.32E-01	−0.17	−0.36 to 0.02	8.50E-02
(0 = less than 1–2 weekly, 1 = more than 1–2 weekly)	0.5	−0.17	−0.56 to 0.22	3.92E-01	−0.65	−1.22 to −0.09	2.32E-02*
	1	−0.19	−0.58 to 0.20	3.42E-01	−0.54	−1.07 to −0.00	4.83E-02*
Employment mother	0.01	0	−0.10 to 0.10	9.67E-01	0.11	0.03–0.19	6.38E-03*
(0 = employed, 1 = unemployed)	0.5	0.05	−0.11 to 0.21	5.42E-01	0.19	−0.01 to 0.40	6.37E-02
	1	0.04	−0.12 to 0.20	5.97E-01	0.19	−0.01 to 0.38	6.17E-02
Father reads	0.01	0.09	−0.15 to 0.32	4.61E-01	0.08	−0.10 to 0.25	3.77E-01
(0 = daily/weekly, 1 = never/monthly)	0.5	−0.11	−0.47 to 0.24	5.36E-01	0.59	0.08–1.09	2.27E-02*
	1	−0.11	−0.47 to 0.26	5.62E-01	0.61	0.13–1.08	1.18E-02*
Mother reads	0.01	−0.03	−0.24 to 0.18	7.64E-01	0.15	0.00–0.31	4.94E-02*
(0 = daily/weekly, 1 = never/monthly)	0.5	0.07	−0.24 to 0.39	6.49E-01	0.6	0.20–1.01	3.42E-03*
	1	0.06	−0.26 to 0.38	7.21E-01	0.63	0.24–1.01	1.31E-03*

* = significant, ** = significant after correcting for multiple testing (≤ 5.81×10^{-4}), Not all thresholds have been included (0.1, 0.2, 0.4, 0.4 have been omitted). Beta, Beta Coefficient; CI, Confidence Interval; MCS, Millennium Cohort Study; PRS, Polygenic Risk Score; rGE, gene–environment correlation.

difference may reflect a cohort effect in environmental risk given that only 70% of mothers were married in MCS compared to over 97% of mothers in NCDS. On the other hand, two of the three indicators of low SES were associated with the genetic risk for MDD across both cohorts, suggesting stability across generations.

Moreover, paternal lack of involvement in child care was associated with the genetic liability to SCZ in NCDS only and may highlight cultural differences in gender responsibilities (Davis & King, 2018) in individuals born in 1958 where the fathers would have been less involved in their offspring's upbringing.

Furthermore, associations between maternal and paternal lack of interest in the child's education and genetic risk for MDD emerged in NCDS only. One interpretation is that with schools having taken on a stronger role in monitoring children's education (Davies & Brember, 2001), differences in parental support today may be less influential than for children born back in 1958. Finally, the genetic risk for MDD was associated with more maternal smoking and less alcohol consumption in MCS but not NCDS. One possible explanation for these associations could be cultural changes in environment risk, which has been re-emphasised by a recent study investigating maternal smoking in pregnancy using the same cohorts as ours (MCS and NCDS) confirming a stronger association between smoking behaviour and social disadvantage in MCS compared to NCDS (Sellers et al., 2020).

Limitations

Although our study features many strengths such as genome-wide data and two large longitudinal samples, our findings have to be considered in light of multiple limitations. First, the revised GWAS results to create PRSs for NCDS excluded several UK cohorts, which may have reduced the chance to detect significant rGEs in NCDS. Second, analyses for some risk factors were underpowered (i.e. father's interest in the child's education at wave 4 in MCS and at ages 7, 11 and 16 in NCDS). Third, there were several significant differences between the original cohorts and the samples included in this analysis. Fourth, whilst the majority of environmental variables were comparable across the two cohorts, the match is not perfect. Fifth, our study included UK cohorts only, and therefore, our results may not be generalisable to other countries. Finally, parental genetic data were not available for NCDS.

Implications

Given that several established psychosocial and behavioural childhood risk factors for psychopathology are associated with the genetic risk for

Table 2 Regression results of rGE in the NCDS sample

	SCZ				MDD		
Environment	PRS Threshold *z*-scored	Beta	95%CI	*p*-Value	Beta	95%CI	*p*-Value
SES	0.01	0.01	−0.02 to 0.03	6.33E-01	−0.06	−0.08 to −0.03	**3.56E-06****
(5 Professional, 4 Managerial/technical,	0.5	0	−0.02 to 0.03	8.34E-01	−0.05	−0.07 to −0.03	**2.53E-05****
3 Skilled, 2 Partly skilled, 1 Unskilled)	1	0	−0.02 to 0.03	8.24E-01	−0.05	−0.07 to −0.03	**4.66E-05****
Finance issues	0.01	0.04	−0.07 to 0.16	4.57E-01	0.17	0.06 to 0.28	1.83E-03*
(0 = no, 1 = yes)	0.5	0.05	−0.06 to 0.16	3.97E-01	0.15	0.05 to 0.26	4.69E-03*
	1	0.04	−0.07 to 0.15	4.78E-01	0.15	0.05 to 0.26	4.40E-03*
Number of rooms	0.01	0	−0.02 to 0.03	8.47E-01	−0.05	−0.07 to −0.02	**2.08E-04****
(continuous)	0.5	0	−0.03 to 0.02	7.79E-01	−0.04	−0.07 to −0.02	**4.44E-04****
	1	0	−0.03 to 0.02	7.41E-01	−0.04	−0.07 to −0.02	7.68E-04*
Tenure	0.01	−0.07	−0.29 to 0.16	5.56E-01	0.47	0.26 to 0.68	**9.88E-06****
(0 = owns, 1 = rents)	0.5	0	−0.21 to 0.21	9.88E-01	0.31	0.11 to 0.52	2.76E-03*
	1	−0.01	−0.22 to 0.20	9.42E-01	0.29	0.08 to 0.49	6.02E-03*
Mother's interest in child's education	0.01	0.05	−0.05 to 0.15	3.13E-01	0.17	0.08 to 0.26	**1.81E-04****
(0 = interested, 1 = uninterested)	0.5	0.05	−0.05 to 0.14	3.32E-01	0.13	0.04 to 0.22	4.61E-03*
	1	0.05	−0.05 to 0.14	3.40E-01	0.12	0.03 to 0.21	7.13E-03*
Father's involvement in child care	0.01	0.19	0.07 to 0.31	1.51E-03*	0.05	−0.05 to 0.16	3.34E-01
(0 = involved, 1 = uninvolved)	0.5	0.21	0.10 to 0.32	**2.50E-04****	0.06	−0.05 to 0.16	2.95E-01
	1	0.2	0.09 to 0.32	**3.00E-04****	0.06	−0.05 to 0.16	2.90E-01
Father's interest in child's education	0.01	0.02	−0.09 to 0.14	6.65E-01	0.2	0.10 to 0.30	**1.21E-04****
(0 = interested, 1 = uninterested)	0.5	0.03	−0.08 to 0.13	5.98E-01	0.19	0.09 to 0.29	**2.04E-04****
	1	0.02	−0.09 − 0.12	7.14E-01	0.18	0.08 to 0.28	**4.73E-04****
Mother walks	0.01	0.04	-0.12 to 0.20	6.20E-01	0.14	−0.01 to 0.29	5.93E-02
(0 = Most weeks/occasionally 1 = Hardly ever)	0.5	0.07	−0.08 to 0.22	3.68E-01	0.2	0.06 to 0.35	7.23E-03*
	1	0.07	−0.09 to 0.22	3.92E-01	0.21	0.06 to 0.36	5.07E-03*
Father walks	0.01	0.02	−0.11 to 0.15	7.56E-01	0.15	0.02 to 0.27	1.96E-02*
(0 = Most weeks/occasionally 1 = Hardly ever)	0.5	0.07	−0.06 to 0.19	3.13E-01	0.1	−0.02 to 0.22	1.10E-01
	1	0.06	−0.07 to 0.19	3.47E-01	0.1	−0.02 to 0.22	9.69E-02
Maternal smoking prior pregnancy	0.01	0.03	−0.03 to 0.09	3.44E-01	0.02	−0.04 to 0.08	4.46E-01
(0 = no, 1 = yes)	0.5	0.03	−0.03 to 0.08	4.05E-01	0.03	−0.02 to 0.09	2.54E-01
	1	0.02	−0.04 to 0.08	4.26E-01	0.03	−0.02 to 0.09	2.69E-01
Maternal smoking during pregnancy	0.01	0.03	−0.04 to 0.09	4.25E-01	0.04	−0.02 to 0.10	1.87E-01
(0 = no, 1 = yes)	0.5	0.01	−0.06 to 0.07	8.27E-01	0.03	−0.03 to 0.09	2.62E-01
	1	0.01	−0.06 to 0.07	8.24E-01	0.03	−0.03 to 0.09	2.66E-01
Parity	0.01	0	−0.03 to 0.03	9.73E-01	0	−0.02 to 0.03	7.59E-01
(continuous)	0.5	−0.01	−0.04 − 0.02	4.89E-01	0.01	-0.02 to 0.03	6.24E-01
	1	−0.01	−0.04 − 0.02	4.15E-01	0.01	-0.02 − 0.03	6.70E-01
Mother's age	0.01	0	-0.03 to 0.03	7.84E-01	−0.04	−0.06 to −0.01	1.05E-02*
(continuous)	0.5	0	−0.03 − 0.02	7.45E-01	-0.02	−0.05 − 0.01	1.74E-01
	1	-0.01	−0.03 − 0.02	6.97E-01	-0.02	−0.05 to 0.01	2.02E-01
Father's age	0.01	0	−0.03 − 0.03	9.83E-01	0	-0.03 to 0.03	9.72E-01
(continuous)	0.5	−0.01	−0.03 − 0.02	7.27E-01	0	-0.02 to 0.03	8.21E-01
	1	−0.01	−0.03 to 0.02	7.22E-01	0	−0.03 to 0.03	8.48E-01

(continued)

Table 2 (continued)

Environment	SCZ				MDD		
	PRS Threshold *z*-scored	Beta	95%CI	*p*-Value	Beta	95%CI	*p*-Value
Gestational period	0.01	−0.01	−0.05 − 0.02	3.79E-01	-0.02	−0.05 to 0.01	2.65E-01
(continuous)	0.5	0	−0.04 to 0.03	7.61E-01	−0.01	−0.04 to 0.02	5.77E-01
	1	0	−0.03 to 0.03	8.40E-01	−0.01	−0.04 to 0.02	4.95E-01
Birth weight	0.01	0	−0.03 − 0.03	8.97E-01	-0.03	−0.06 to −0.00	3.09E-02
(continuous)	0.5	0	−0.03 to 0.03	9.49E-01	−0.02	−0.05 to 0.01	1.96E-01
	1	0	−0.03 to 0.03	9.68E-01	−0.02	−0.05 to 0.01	1.89E-01
Marital status	0.01	0.06	−0.12 to 0.24	4.97E-01	0.08	−0.09 to 0.25	3.39E-01
(0 = married, stable union, 1 = divorced, separated)	0.5	−0.06	−0.23 to 0.11	4.84E-01	0.05	−0.12 to 0.21	5.59E-01
	1	−0.07	−0.24 to 0.10	4.26E-01	0.04	−0.12 to 0.21	6.19E-01
Housing issues	0.01	0.13	−0.01 to 0.27	6.87E-02	0.15	0.02 to 0.28	1.89E-02
(0 = no, 1 = yes)	0.5	0.16	0.03 to 0.29	1.45E-02*	0.09	−0.04 to 0.21	1.65E-01
	1	0.17	0.04 to 0.30	1.04E-02*	0.08	−0.05 to 0.20	2.15E-01
Family alcohol issues	0.01	0.23	−0.14 to 0.60	2.17E-01	0.13	−0.21 to 0.48	4.39E-01
(0 = no, 1 = yes)	0.5	0.06	−0.29 to 0.42	7.30E-01	0.22	−0.12 to 0.57	1.98E-01
	1	0.03	−0.32 to 0.38	8.63E-01	0.23	−0.11 to 0.58	1.78E-01
Domestic tension	0.01	0.21	0.05 to 0.37	8.71E-03*	0.13	−0.01 to 0.28	7.03E-02
(0 = no, 1 = yes)	0.5	0.18	0.03 to 0.33	2.04E-02*	0.12	−0.02 to 0.27	8.76E-02
	1	0.17	0.02 to 0.32	2.61E-02*	0.13	−0.02 to 0.27	8.26E-02
Employment father	0.01	0.01	−0.24 to 0.25	9.64E-01	−0.02	−0.24 to 0.21	8.77E-01
(0 = employed 1 = unemployed)	0.5	0.07	−0.16 to 0.30	5.78E-01	0.16	−0.06 to 0.38	1.61E-01
	1	0.06	−0.17 to 0.29	5.91E-01	0.17	−0.06 to 0.39	1.45E-01
Father reads	0.01	0.08	0.00 to 0.15	3.89E-02*	0.01	−0.06 to 0.08	7.68E-01
(0 = weekly/occasionally 1 = Hardly ever)	0.5	0.08	0.01 to 0.15	2.17E-02*	0.01	−0.06 to 0.07	8.44E-01
	1	0.09	0.02 to 0.16	1.67E-02*	0	−0.06 to 0.07	8.98E-01
Mother reads	0.01	−0.02	−0.11 − 0.08	7.48E-01	0.01	-0.08 to 0.09	8.75E-01
(0 = weekly/occasionally 1 = Hardly ever)	0.5	0.07	−0.01 to 0.16	9.37E-02	0.03	−0.06 to 0.11	5.53E-01
	1	0.08	−0.01 to 0.16	8.01E-02	0.02	−0.06 to 0.10	6.17E-01
Free school meals	0.01	0.07	−0.11 to 0.26	4.42E-01	0.09	−0.09 to 0.26	3.31E-01
(0 = no, 1 = yes)	0.5	0.09	−0.09 to 0.26	3.34E-01	0.12	−0.05 to 0.29	1.80E-01
	1	0.08	−0.09 to 0.26	3.51E-01	0.12	−0.05 to 0.29	1.69E-01

* = significant, ** = significant after correcting for multiple testing (≤ 5.81×10^{-4}), Not all thresholds have been included (0.1, 0.2, 0.4, 0.4 have been omitted). Beta, Beta Coefficient; CI, Confidence Interval; NCDS, 1958 National Child Development Study; PRS, Polygenic Risk Score; rGE, gene–environment correlation.

psychiatric disorders in our data, findings may suggest that interventions that target such environmental risk factors may have little effect on individuals with at substantial genetic risk (Wagner, Li, Liu, & Guo, 2013). However, it is important to consider that PRSs for psychopathology usually only explain a small portion of the variance, for example, 2% in the case of MDD (Lewis & Vassos, 2020), and it is likely that the majority of children in a community sample with an elevated genetic risk for MDD will grow up into healthy adults. Hence, our results do not suggest that targeting environmental risk will necessarily be inefficient for children with a genetic risk for psychiatric disorders. However, given the observation of passive rGE for several environmental risk factors, it may be helpful to prioritise systemic approaches that focus on both children and their parents.

Conclusions

We identified several correlations between known environmental risk factors in childhood and genetic risk for SCZ and MDD in two independent community samples. Gene–environment correlation was more pronounced for PRS of MDD and less so for SCZ. More than half of detected correlations (in MCS) were confounded by the parental genetic risk and therefore represent passive rGE, whereby parents shape the environment in addition to passing on their genes. Importantly, there was little overlap between rGEs of SCZ and MDD, suggesting that rGE differs between disorders. Furthermore, findings also suggest that whilst some rGE are stable across generations (e.g. low SES), others likely change due to societal and cultural changes such as smoking behaviour, the proportion of single-parent families and father's involvement in child care. In sum, findings confirm the complex relationship between environmental and genetic risk for psychiatric disorders and emphasise the importance of considering gene–environment interplay regarding the role of long-established environmental risk factors in childhood.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. Environmental Risk Factors for SCZ and MDD.
Appendix S2. Genetic QC.
Appendix S3. Coded variables.
Appendix S4. Descriptive statistics.
Appendix S5. Full SCZ and MDD results.

Acknowledgements

The authors would like to dedicate this article to their dear friend, colleague and supervisor Dr Robert Keers who sadly passed away during the reported work. The authors would like to acknowledge his significant contribution to the conceptual design of the research, supervision of the genetic QC and support of the statistical analyses. He is deeply missed and fondly remembered.

The authors declare that this project received funding from GlaxoSmithKline R&D. The funder was not involved in the study design, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

D.M.H. is supported by a Sir Henry Wellcome Postdoctoral Fellowship (Reference 213674/Z/18/Z) and a 2018 NARSAD Young Investigator Grant from the Brain & Behaviour Research Foundation (Ref: 27404). P.B.M. wishes to acknowledge support from the NIHR Barts Biomedical Research Centre.

Data governance was provided by the METADAC data access committee, funded by ESRC, Wellcome, and MRC. (2015–2018: Grant Number MR/N01104X/1 2018–2020: Grant Number ES/S008349/1).

This work made use of data and samples generated by the 1958 Birth Cohort (NCDS), which is managed by the Centre for Longitudinal Studies at the UCL Institute of Education, funded by the Economic and Social Research Council (grant number ES/M001660/1). Access to these resources was enabled via the Wellcome Trust & MRC: 58FORWARDS grant [108439/Z/15/Z] (The 1958 Birth Cohort: Fostering new Opportunities for Research via Wider Access to Reliable Data and Samples). Before 2015 biomedical resources were maintained under the Wellcome Trust and Medical Research Council 58READIE Project (grant numbers WT095219MA and G1001799). Genotyping was undertaken as part of the Wellcome Trust Case–Control Consortium (WTCCC) under Wellcome Trust award 076113, and a full list of the investigators who contributed to the generation of the data is available at www.wtccc.org.uk.

This study also made use of data from the Millennium Cohort Study.

The authors are grateful to the Centre for Longitudinal Studies (CLS), UCL Social Research Institute, for the use of these data and to the UK Data Service for making them available. However, neither CLS nor the UK Data Service bear any responsibility for the analysis or interpretation of these data.

S.M-N. is a part-time employee at GlaxoSmithKline R&D. The author did not receive any financial incentive for this research. The remaining authors have declared that they have no competing or potential conflicts of interest.

Correspondence

Michael Pluess, Department of Biological and Experimental Psychology, Queen Mary, University of London, Mile End Road, London E1 4NS, UK; Email: m.pluess@qmul.ac.uk

Key points

- Established environmental and psychosocial risk factors for schizophrenia and depression may be associated with genetic risk for psychopathology in children.
- We investigated correlations between polygenic scores (schizophrenia and depression) and established environmental risk factors in two large British cohort studies.
- Findings suggest that the genetic risk for schizophrenia in children is correlated with parents being divorced, separated or widowed and lack of father's involvement in child care, whereas indicators of low socioeconomic status and adverse parental behavioural and psychosocial factors such as lack of interest in the child's education were associated with the genetic risk for depression.
- According to the analyses of genetic data from both children and parents, several of the detected associations were confounded by the parents' genetic risk with important implications for the prevention and treatment of complex psychopathologies.

References

Abel, K.M., Wicks, S., Susser, E.S., Dalman, C., Pedersen, M.G., Mortensen, P.B., & Webb, R.T. (2010). Birth weight, schizophrenia, and adult mental disorder: Is risk confined to the smallest babies? *Archives of General Psychiatry*, *67*, 923–930.

Alati, R., Lawlor, D.A., Mamun, A.A., Williams, G.M., Najman, J.M., O'Callaghan, M., & Bor, W. (2007). Is there a fetal origin of depression? Evidence from the mater university study of pregnancy and its outcomes. *American Journal of Epidemiology*, *165*, 575–582.

Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., … Consortium, G.P. (2015). A global reference for human genetic variation. *Nature*, *526*, 68–74.

Bakken Stovner, E. (2017). snpflip. Available from: https://github.com/biocore-ntnu/snpflip. [Retrieved November 03, 2019]

Bann, D., Johnson, W., Li, L., Kuh, D., & Hardy, R. (2018). Socioeconomic inequalities in childhood and adolescent body-mass index, weight, and height from 1953 to 2015: An analysis of four longitudinal, observational, British birth cohort studies. *The Lancet Public Health*, *3*, e194–e203.

Barrett, J.C., Clayton, D.G., Concannon, P., Akolkar, B., Cooper, J.D., Erlich, H.A., … Consortium, T.D.G. (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature Genetics*, *41*, 703–707.

Brown, A.S., & Derkits, E.J. (2010). Prenatal infection and schizophrenia: A review of epidemiologic and translational studies. *The American Journal of Psychiatry*, *167*, 261–280.

Burns, J.K., Tomita, A., & Kapadia, A.S. (2014). Income inequality and schizophrenia: Increased schizophrenia incidence in countries with high levels of income inequality. *The International Journal of Social Psychiatry*, *60*, 185–196.

Cannon, T.D., Rosso, I.M., Hollister, J.M., Bearden, C.E., Sanchez, L.E., & Hadley, T. (2000). A prospective cohort study of genetic and perinatal influences in the etiology of schizophrenia. *Schizophrenia Bulletin*, *26*, 351–366.

Centre for Longitudinal Studies. (2014). *National Child Development Study – Ethical review and consent*. London, UK: Centre for Longitudinal Studies. Available from: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/NCDS-Ethical-review-and-Consent-2014.pdf [Retrieved 12 May 2022]

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., & Lee, J.J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*, *4*, 7.

Chiu, T.F., Yu, T.M., Chuang, Y.W., Sun, K.T., Li, C.Y., Su, Y.C., & Kao, C.H. (2019). Sequential risk of depression in children born prematurely: A nationwide population- based analysis. *Journal of Affective Disorders*, *243*, 42–47.

Cohen, A.K., Nussbaum, J., Weintraub, M.L.R., Nichols, C.R., & Yen, I.H. (2020). Association of adult depression with educational attainment, aspirations, and expectations. *Preventing Chronic Disease*, *17*, E94.

Coleman, J.R., Euesden, J., Patel, H., Folarin, A.A., Newhouse, S., & Breen, G. (2016). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in Functional Genomics*, *15*, 298–304.

Connelly, R., & Platt, L. (2014). Cohort profile: UK millennium cohort study (MCS). *International Journal of Epidemiology*, *43*, 1719–1725.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., … Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, *10*, giab008. https://doi.org/10.1093/gigascience/giab008

Davies, J., & Brember, I. (2001). A decade of change? Monitoring reading and mathematical attainment in year 6 over the first ten years of the education reform act. *Research in Education*, *65*, 31–40.

Davis, A., & King, L. (2018). Gendered perspectives on Men's changing familial roles in postwar England. *Gender & History*, *30*, 70–92.

Dudbridge, F., & Newcombe, P.J. (2015). Accuracy of gene scores when pruning markers by linkage disequilibrium. *Human Heredity*, *80*, 178–186.

Ensink, J.B.M., de Moor, M.H.M., Zafarmand, M.H., de Laat, S., Uitterlinden, A., Vrijkotte, T.G.M., … Middeldorp, C.M. (2020). Maternal environmental risk factors and the development of internalizing and externalizing problems in childhood: The complex role of genetic factors. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *183*, 17–25.

Euesden, J., Lewis, C.M., & O'Reilly, P.F. (2015). PRSice: Polygenic risk score software. *Bioinformatics*, *31*, 1466–1468.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Fitzsimons, E., Moulton, V., Hemani, D.A., Ho, K., Hughes, G., Neaves, S., … Ring, S. (2020). Collection of DNA samples and genetic data at scale in the UKmillennium cohort study. In *CLS working paper number 2020/7*. London: UCL Centre for Longitudinal Studies.

Freeman, A., Tyrovolas, S., Koyanagi, A., Chatterji, S., Leonardi, M., Ayuso-Mateos, J.L., … Haro, J.M. (2016). The role of socio-economic status in depression: Results from the COURAGE (aging survey in Europe). *BMC Public Health*, *16*, 1098.

Glowinski, A.L., Madden, P.A., Bucholz, K.K., Lynskey, M.T., & Heath, A.C. (2003). Genetic epidemiology of self-reported lifetime DSM-IV major depressive disorder in a population-based twin sample of female adolescents. *Journal of Child Psychology and Psychiatry*, *44*, 988–996.

Harrison, G., Gunnell, D., Glazebrook, C., Page, K., & Kwiecinski, R. (2001). Association between schizophrenia and social inequality at birth: Case-control study. *The British Journal of Psychiatry*, *179*, 346–350.

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T.M., … Glenthøj, B. (2018). Heritability of schizophrenia and schizophrenia Spectrum based on the Nationwide Danish twin register. *Biological Psychiatry*, *83*, 492–498.

Jaffee, S.R., & Price, T.S. (2007). Gene-environment correlations: A review of the evidence and implications for prevention of mental illness. *Molecular Psychiatry*, *12*, 432–442.

Jaffee, S.R., & Price, T.S. (2008). Genotype-environment correlations: Implications for determining the relationship between environmental exposures and psychiatric illness. *Psychiatry*, *7*, 496–499.

Jaffee, S.R., & Price, T.S. (2012). The implications of genotype-environment correlation for establishing causal processes in psychopathology. *Development and Psychopathology*, *24*, 1253–1264.

Jones, P.B., Rantakallio, P., Hartikainen, A.L., Isohanni, M., & Sipila, P. (1998). Schizophrenia as a long-term outcome of pregnancy, delivery, and perinatal complications: A 28-year follow-up of the 1966 North Finland general population birth cohort. *The American Journal of Psychiatry*, *155*, 355–364.

Jones, H.J., Stergiakouli, E., Tansey, K.E., Hubbard, L., Heron, J., Cannon, M., … Zammit, S. (2016). Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry*, *73*, 221–228.

Kahn, R.S., Sommer, I.E., Murray, R.M., Meyer-Lindenberg, A., Weinberger, D.R., Cannon, T.D., … Insel, T.R. (2015). Schizophrenia. *Nature Reviews. Disease Primers*, *1*, 15067.

Keefe, R.S., Eesley, C.E., & Poe, M.P. (2005). Defining a cognitive function decrement in schizophrenia. *Biological Psychiatry*, *57*, 688–691.

Kendler, K.S., & Karkowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: Genetic control of exposure to the environment? *Psychological Medicine*, *27*, 539–547.

Kendler, K.S., & Prescott, C.A. (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry*, *56*, 39–44.

Krapohl, E., Hannigan, L.J., Pingault, J.B., Patel, H., Kadeva, N., Curtis, C., … Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 11727–11732.

Kwong, A.S.F., Morris, T.T., Pearson, R.M., Timpson, N.J., Rice, F., Stergiakouli, E., & Tilling, K. (2021). Polygenic risk for depression, anxiety and neuroticism are associated with the severity and rate of change in depressive symptoms across adolescence. *Journal of Child Psychology and Psychiatry*, *62*, 1462–1474.

Lewis, C.M., & Vassos, E. (2020). Polygenic risk scores: From research tools to clinical instruments. *Genome Medicine*, *12*, 44.

Nasir, M., & Bloch, M.H. (2021). Editorial: Money cannot buy happiness – but can it prevent depression? A commentary on Su et al. *Journal of Child Psychology and Psychiatry*, *62*, 1047–1049.

Nivard, M.G., Gage, S.H., Hottenga, J.J., van Beijsterveldt, C.E.M., Abdellaoui, A., Bartels, M., … Middeldorp, C.M. (2017). Genetic overlap between schizophrenia and developmental psychopathology: Longitudinal and multivariate polygenic risk prediction of common psychiatric traits during development. *Schizophrenia Bulletin*, *43*, 1197–1207.

Patterson, N., Price, A.L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, *2*, e190.

Plomin, R., DeFries, J.C., & Loehlin, J.C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309–322.

Power, C., & Elliott, J. (2006). Cohort profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*, *35*, 34–41.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*, 904–909.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., … Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575.

Rice, F. (2010). Genetics of childhood and adolescent depression: Insights into etiological heterogeneity and challenges for future genomic research. *Genome Medicine*, *2*, 68.

Ritchie, S. (2014). liftOverPlink. In *GitHub*. Available from: https://github.com/sritchie73/liftOverPlink. [Retrieved November 03, 2019]

Sacker, A., Done, D.J., Crow, T.J., & Golding, J. (1995). Antecedents of schizophrenia and affective illness. Obstetric complications. *The British Journal of Psychiatry*, *166*, 734–741.

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*, 421–427.

Sellers, R., Warne, N., Rice, F., Langley, K., Maughan, B., Pickles, A., … Collishaw, S. (2020). Using a cross-cohort comparison design to test the role of maternal smoking in pregnancy in child mental health and learning: Evidence from two UKcohorts born four decades apart. *International Journal of Epidemiology*, *49*, 390–399.

Shepherd, P., & Gilbert, E. (2019). *Millennium cohort study – Ethical review and consent* (In edn). London: Centre for Longitudinal Studies.

Staatz, C.B., Kelly, Y., Lacey, R.E., & Hardy, R. (2021). Area-level and family-level socioeconomic position and body composition trajectories: Longitudinal analysis of the UKmillennium cohort study. *The Lancet Public Health*, *6*, e598–e607.

StataCorp. (2011). *Stata statistical software: Release 12*. College Station, TX: StataCorp LP.

Sullivan, P.F., Kendler, K.S., & Neale, M.C. (2003). Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, *60*, 1187–1192.

Sullivan, P.F., Neale, M.C., & Kendler, K.S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, *157*, 1552–1562.

Wagner, B., Li, J., Liu, H., & Guo, G. (2013). Gene-environment correlation: Difficulties and a natural experiment-based strategy. *American Journal of Public Health*, *103*(Suppl 1), S167–S173.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, *447*, 661–678.

Wellcome Trust Case Control Consortium 2. (n.d.). Available from: www.wtccc.org.uk/ccc2/. [Retrieved 23 September 2021]

Wertz, J., Belsky, J., Moffitt, T.E., Belsky, D.W., Harrington, H., Avinun, R., ... Caspi, A. (2019). Genetics of nurture: A test of the hypothesis that parents' genetics predict their observed caregiving. *Developmental Psychology*, *55*, 1461–1472.

Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., ... Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*, 668–681.

Accepted for publication: 20 May 2022