

Effect of polygenic risk scores on depression in childhood trauma

Wouter J. Peyrot, Yuri Milaneschi, Abdel Abdellaoui, Patrick F. Sullivan, Jouke J. Hottenga, Dorret I. Boomsma and Brenda W. J. H. Penninx

Background

Research on gene \times environment interaction in major depressive disorder (MDD) has thus far primarily focused on candidate genes, although genetic effects are known to be polygenic.

Aims

To test whether the effect of polygenic risk scores on MDD is moderated by childhood trauma.

Method

The study sample consisted of 1645 participants with a DSM-IV diagnosis of MDD and 340 screened controls from The Netherlands. Chronic or remitted episodes (severe MDD) were present in 956 participants. The occurrence of childhood trauma was assessed with the Childhood Trauma Interview and the polygenic risk scores were based on genome-wide meta-analysis results from the Psychiatric Genomics Consortium.

Results

The polygenic risk scores and childhood trauma

independently affected MDD risk, and evidence was found for interaction as departure from both multiplicativity and additivity, indicating that the effect of polygenic risk scores on depression is increased in the presence of childhood trauma. The interaction effects were similar in predicting all MDD risk and severe MDD risk, and explained a proportion of variation in MDD risk comparable to the polygenic risk scores themselves.

Conclusions

The interaction effect found between polygenic risk scores and childhood trauma implies that (1) studies on direct genetic effect on MDD gain power by focusing on individuals exposed to childhood trauma, and that (2) individuals with both high polygenic risk scores and exposure to childhood trauma are particularly at risk for developing MDD.

Declaration of interest

None.

Research on gene \times environment (G \times E) interaction in major depressive disorder (MDD) aims to understand the heterogeneity of environmental and genetic risk factors, but has thus far primarily focused on candidate genes with inconclusive findings.^{1,2} On the one hand, research on G \times E interaction could select individuals with increased vulnerability for environmental factors based on their genetic make-up. Alternatively, research on G \times E interaction could select environmental conditions that lead to increased expression of genetic effects. Insights into G \times E interaction is therefore of general importance for psychiatric research and contributes to the understanding of MDD's complex aetiology.

Thus far, G \times E interaction has primarily been tested for candidate genes such as the serotonin transporter gene (5-HTTLPR), for which opposing results were found in very similar single studies,^{3,4} as well as in meta-analyses.^{5,6} Several environmental factors have been analysed in this respect, and among the most important factors is childhood trauma, which has a strong impact on MDD risk.^{7–10} Nevertheless, although some consistent evidence for interaction between childhood trauma and 5-HTTLPR was found, these G \times E findings remain controversial.¹ The progress from a candidate gene to a hypothesis-free genome-wide approach is hampered by lack of statistical power and inconsistent assessment of environmental stressors across genome-wide association study (GWAS) cohorts.

Research on main genetic effects, on the other hand, has indicated that the risk of MDD is not merely increased by the effect of one or a few single nucleotide polymorphisms (SNPs), but by polygenic variation.^{11,12} One of the methods applied to point at these polygenic effects, first introduced for schizophrenia,¹³ uses polygenic risk scores and was later applied to MDD.¹¹ The polygenic risk scores are obtained after carrying out a GWAS in a discovery sample and then taking SNPs up to a certain threshold

of significance, or even all SNPs, to predict MDD in an independent target sample. The contribution of these large numbers of SNPs are weighted by their effect size in a GWAS or meta-analysis. The effect of polygenic risk scores on MDD was repeatedly confirmed and explains up to 1–2% of variation.^{11,14,15}

Even though this has not yet been studied for MDD, it is likely that causal genetic variants for MDD are located throughout all of the genome, as has been found for other complex traits such as height and body mass index.¹⁶ Also, SNPs contributing to pleiotropy between schizophrenia and bipolar disorder are found dispersed throughout the genome.¹⁷ The finding that the risk of MDD is increased by polygenic variation suggests that research of interaction effects should also focus on polygenic information. With an expected abundance of causal variants for MDD, environmental conditions that increase genetic effects are more likely to be found when polygenic information is taken into account. When environmental conditions that increase genetic effects are found, individuals exposed to these conditions can be selected for future research to study the impact of single loci on MDD with increased power. Nevertheless, to the best of our knowledge, no research on G \times E interaction in MDD has focused on polygenic information thus far.

The current study focused on polygenic information to test for G \times E interaction in MDD, and examined whether polygenic risk scores interact with the presence of childhood trauma in a large and well-characterised sample from The Netherlands.

Method

Participants

The sample consisted of participants from the Netherlands Study of Depression and Anxiety (NESDA), which is an ongoing

longitudinal cohort study of depressive and anxiety disorders, with individuals recruited from mental healthcare settings, general practices and the general population in the period from 2004 to 2007.¹⁸ Participants with MDD in their lifetime ($n = 1645$) were diagnosed in a face-to-face interview with a trained clinical staff member following the DSM-IV-based Composite International Diagnostic Interview (CIDI, version 2.1). Over half of these participants ($n = 956$) had severe MDD with remitted (more than one) episodes and/or chronic (longer than 2 years of) complaints, as assessed with the life chart, a calendar approach to calculate the percentage of time symptoms were present during 4 years prior to baseline and 2 years following baseline.¹⁹ Controls ($n = 340$) were screened in a similar face-to-face CIDI interview and had no diagnosis of a depressive, dysthymic, anxiety or other psychiatric disorder in their lifetime. Participants were of North European ancestry and were excluded when they were not fluent in speaking Dutch or when they had another primary diagnosis, such as a psychotic, obsessive-compulsive, bipolar or severe substance use disorder. The NESDA study was approved by the institutional review board and all participants provided written informed consent.

Childhood trauma

Childhood trauma was assessed in a face-to-face interview with a trained clinical staff member with the Childhood Trauma Interview (CTI) from The Netherlands Mental Health Survey and Incidence Study.⁷ The CTI assesses the domains of emotional neglect, psychological abuse, physical abuse and sexual abuse before the age of 16, and yields a score ranging 0–8 by adding the frequencies of occurrence (0, absent; 1, once or sometimes; 2, regularly, often or very often). In the CTI the four domains are assessed by asking whether the traumatic event occurred (yes or no), and a subsequent question asking how often the event occurred. In the first question the traumatic events were specified as follows: emotional neglect as the lack of parental attention or support and ignorance of one's problems and experiences; psychological abuse as being verbally abused, undeserved punishment, subordinated to siblings and being blackmailed; physical abuse as being kicked or hit with hands or an object, beaten up or physical abuse in any other way; and sexual abuse as being sexually approached against one's will, meaning being touched or having to touch someone in a sexual way. The CTI is a well-established instrument – measurements of childhood trauma show a strong impact on depressive and anxiety disorders^{7,20} as well as on structural and functional brain abnormalities.^{21,22} The CTI also shows strong content validity when compared with the Childhood Trauma Questionnaire (CTQ),²³ with Spearman's rho correlation of 0.69 ($P < 0.001$) in a subset of NESDA with both the CTI and CTQ assessed at different time points.

Genotyping and quality control

Methods for blood sampling and DNA extraction have been described previously.²⁴ The manufacturer's protocol was followed to genotype the autosomal SNPs on the Affymetrix 6.0 Human Genome-Wide SNP Array. With quality control, SNPs were excluded that: had probes that mapped badly against NCBI Build 37/UCSC hg19; had a minor allele frequency smaller than 1%; had a missing rate greater than 5%; deviated from Hardy-Weinberg equilibrium with a P -value smaller than 0.001, thus leaving 498 592 SNPs to analyse. Participants were excluded when: they showed a contrast QC < 0.4 (CQC, a quality metric from Affymetrix representing how well allele intensities separate into

clusters); fell outside of the main cluster of a principal component reflecting a batch effect;²⁵ had a missing rate greater than 5%; had excess genome-wide heterozygosity or inbreeding levels ($F < -0.10$ or > 0.10); had genotypes with inconsistencies regarding reported gender; or had non-European/non-Dutch ancestry as indicated with principal component analysis.²⁵

Polygenic risk scores

The polygenic risk scores were created based on the results from a large meta-analysis from the Psychiatric Genomics Consortium (PGC)¹⁵ excluding participants from the Dutch GWAS cohort²⁶ that included NESDA participants (thus yielding 7544 cases and 7754 controls in the discovery set). Risk scores were obtained following the method described by Purcell and colleagues¹³ with the PLINK software run on Linux.²⁷ From the meta-analysis, SNPs were selected that had an imputation INFO score > 0.9 and minor allele frequency > 0.02 , and low linkage disequilibrium to each other ($r^2 < 0.25$ within 500 kb window, filtering for significance; PLINK-command `-clump-p1 1 -clump-p2 1 -clump-r2 0.25 -clump-kb 500`). The meta-analysis results of SNPs up to eight P -value thresholds (0.001; 0.01; 0.05; 0.1; 0.2; 0.3; 0.4; and 0.5) were selected to compute the polygenic risk scores in our sample; the numbers of SNPs thus included were 150, 1209, 5028, 8905, 16 081, 22 355, 28 018 and 32 870 respectively. The polygenic risk scores were standardised to a mean of zero and standard deviation of one to aid interpretation of results.

Statistical analyses

Participants with MDD were compared with controls with respect to age, gender, and their childhood trauma score (range 0–8) with t -tests for continuous variables and chi-squared tests for binary variables. The effect of polygenic risk scores on the childhood trauma score (i.e. gene-environment correlation) was tested with linear regression, because such an effect could potentially bias tests for interaction.²⁸ Two binary MDD outcomes were analysed as dependent variables: all participants with MDD *v.* controls (all MDD risk) and participants with severe MDD *v.* controls (severe MDD risk). The direct effects of polygenic risk scores (model 1) and the childhood trauma score (model 2) on MDD risks were assessed in separate logistic regression models. Subsequently, tests for interaction were performed with logistic regression to test for interaction as departure from multiplicativity (model 3) and, second, with analyses of relative excess risks due to interaction (RERI, model 4) to test for interaction as departure from additivity. The RERIs were computed with the method described by Knol and colleagues, as $RERI = e^{\beta_{CT} + \beta_{PRS} + \beta_{PRS \times CT}} - e^{\beta_{CT}} - e^{\beta_{PRS}} + 1$.²⁹ The RERI's 95% confidence intervals were computed with bootstrapping with 10 000 iterations. The difference between interaction as departure from additivity and interaction as departure from multiplicativity is that the first represents a situation where the combined effect is larger than the sum of the individual effects of the polygenic risk score and childhood trauma, whereas the latter represents a situation where the combined effect is larger than the product of the individual effects. It has been argued that interaction as departure from additivity is more in line with biological interaction.²⁹

Nagelkerke's R^2 were estimated to assess what proportion of variation in all MDD risk was explained by the polygenic risk scores (PRS) and childhood trauma independently, as well as their interaction PRS \times childhood trauma. Therefore, several R^2 estimates were compared: between the model with only covariates and the model additionally including childhood trauma (R^2 of childhood trauma); between the model with covariates and

childhood trauma and the model additionally including polygenic risk scores (R^2 of PRS); and between the model with covariates, childhood trauma and polygenic risk scores and the model additionally including the PRS \times childhood trauma interaction (R^2 of PRS \times childhood trauma). Nagelkerke's R^2 may, however, be biased by a sample's ascertainment when a disproportionate number of cases is selected from the population.³⁰ Therefore, we also computed an alternative R^2 measure for the polygenic risk scores, which was recently proposed by Lee and colleagues.³⁰ This R^2 measure is based on the liability scale, directly comparable to the heritability, and robust against ascertainment bias.³⁰ Lee's R^2 estimates in our sample were based on a Dutch lifetime prevalence of MDD of 18.7%.³¹

All analyses were corrected for age, gender and three ancestry-informative principal components to take possible population stratification into account, and the tests for interaction (models 3 and 4) included polygenic risk scores and the childhood trauma score as additional covariates. Effects were considered significant when P -values were <0.05 or when RERI 95% confidence intervals did not contain zero. All analyses were performed in R for Windows.³²

Results

Participants with MDD ($n = 1645$) had a mean age comparable to that of the 340 healthy controls (42.2 years (s.d. = 2.5) and 43.3 years (s.d. = 14.5) respectively, $P = 0.172$), and were slightly more often female (68% and 57% respectively, $P < 0.001$). The mean childhood trauma score was 1.75 (s.d. = 2.17, range 0–8), and mean scores of the four childhood trauma domains (range 0–2) were 0.76 (0.95) for emotional neglect (EN), 0.50 (0.84) for psychological abuse (PsA), 0.22 (0.57) for physical abuse (PhA), and 0.24 (0.52) for sexual abuse (SA). The scores of the domains were all correlated with each other with Pearson correlation coefficients of 0.61 for EN–PsA, 0.40 for EN–PhA, 0.24 for EN–SA, 0.55 for PsA–PhA, 0.23 for PsA–SA, and 0.26 for PhA–SA (all $P < 0.001$). Childhood trauma occurred more often in participants with MDD than in healthy controls, with mean childhood trauma main scores of 1.99 (s.d. = 2.24) and 0.56 (s.d. = 1.29) respectively ($P < 0.001$). None of the polygenic risk scores had an effect on childhood trauma, with beta-estimates around zero and all P -values well over 0.05, thus excluding gene–environment correlation and its potential bias on interaction tests (Table 1).

The polygenic risk scores significantly predicted MDD risk (model 1), with slightly larger but comparable effects in predicting severe MDD risk compared with predicting all MDD risk (Table 2). The polygenic risk scores based on five of the eight studied thresholds were predictive in all MDD risk (thresholds 0.05; 0.1; 0.2; 0.3; 0.4) and the polygenic risk scores based on six thresholds were predictive in severe MDD risk (thresholds 0.05; 0.1; 0.2; 0.3; 0.4; 0.5). The score based on threshold $P < 0.05$ had the largest effect on all MDD risk, with an odds ratio (OR) of 1.22 per standard deviation increase of the polygenic risk score ($P = 0.001$). The presence of childhood trauma also predicted MDD risk (model 2), again with slightly larger but comparable effects in predicting severe MDD risk compared with all MDD risk (with ORs of 1.64 ($P < 0.001$) and 1.69 ($P < 0.001$) respectively, per childhood trauma score unit increase (range 0–8); Table 2). Evidence was then found for interaction as departure from both multiplicativity (model 3, ORs > 1) and additivity (model 4, RERIs > 0), indicating that the effect of polygenic risk scores on MDD is increased in the presence of childhood trauma (Table 2). The largest interaction effect in predicting all MDD was found

Table 1 Effect of polygenic risk scores on childhood trauma^a

Polygenic risk score thresholds	Beta	P
$P < 0.001$	< 0.01	0.991
$P < 0.01$	-0.01	0.769
$P < 0.05$	0.02	0.733
$P < 0.1$	0.01	0.847
$P < 0.2$	-0.01	0.883
$P < 0.3$	-0.02	0.754
$P < 0.4$	-0.01	0.904
$P < 0.5$	0.01	0.907

a. Effects of polygenic risk scores on childhood trauma (i.e. gene–environment correlation) were estimated with linear regression including three principal components, age and gender as covariates.

for the polygenic risk score based on threshold $P < 0.1$ with an OR of 1.15 ($P = 0.005$); the largest interaction effect in predicting severe MDD was found for the polygenic risk score based on threshold $P < 0.3$ with an OR of 1.16 ($P = 0.005$). These two interaction effects were visualised for their departure of multiplicativity by displaying the direct effects of the polygenic risk scores for three childhood trauma levels, with childhood trauma scores of 0–1; 2–3 and 4–8 respectively (Fig. 1). Figure 1 shows that the polygenic risk scores have limited impact in predicting MDD risk in individuals with no/low exposure to childhood trauma, but large impact in individuals with high exposure to childhood trauma.

The impact of the four separate childhood trauma domains on the interaction effects were compared by conducting analyses of each domain separately in predicting all MDD risk. The estimates of interaction thus found were in the same direction for all domains (OR > 1), but appeared more significant for the domains of emotional neglect and psychological abuse, than for the domains of physical abuse and sexual abuse (Table 3). This difference in significance is possibly due to the lower frequency of occurrence of physical abuse and sexual abuse.

Most variation in all MDD risk was explained by childhood trauma (~13%), but the proportions explained by the polygenic risk scores (in addition to the variation explained by childhood trauma) and their interaction effects (in addition to the variation explained by childhood trauma and the polygenic risk score) were of comparable magnitude (~0.5%, Table 4). Note that Lee's R^2 estimates were comparable to Nagelkerke's R^2 estimates for the polygenic risk scores, which indicates that ascertainment bias did not largely impact our results (Table 4).

Discussion

Main findings

This is the first study that focuses on polygenic risk scores to test for G \times E interaction in MDD. Within our sample we found increased effects of polygenic risk scores on MDD in the presence of childhood trauma, with evidence for interaction as departure from both multiplicativity and additivity. These interaction effects were comparable in predicting all MDD risk and severe (chronic or recurrent) MDD risk, although effects were slightly larger in the latter. The interaction effects were driven by all of the four domains included in the childhood trauma measure (emotional neglect, psychological abuse, physical abuse and sexual abuse). We found that the proportion of variation in all MDD risk explained by the interaction effects was comparable to the proportion explained by the polygenic risk scores (both ~0.5%).

Table 2 Interaction between polygenic risk score (PRS) and childhood trauma in predicting major depressive disorder risk and direct effects of PRSs and childhood trauma

PRS thresholds	Direct effects ^a				PRS × childhood trauma interaction ^b			
	PRS (model 1)		Childhood trauma (model 2)		Multiplicative (model 3)		Additive (model 4)	
	OR	P	OR	P	OR	P	RERI	95% CI
All major depressive disorder^c								
<i>P</i> < 0.001	1.01	0.808	1.64	< 0.001	1.06	0.288	0.08	−0.08 to 0.25
<i>P</i> < 0.01	1.12	0.059	1.64	< 0.001	1.09	0.080	0.21	0.04 to 0.47
<i>P</i> < 0.05	1.22	0.001	1.64	< 0.001	1.14	0.008	0.37	0.14 to 0.71
<i>P</i> < 0.1	1.18	0.005	1.64	< 0.001	1.15	0.005	0.34	0.13 to 0.64
<i>P</i> < 0.2	1.15	0.021	1.64	< 0.001	1.12	0.014	0.29	0.10 to 0.56
<i>P</i> < 0.3	1.13	0.037	1.64	< 0.001	1.14	0.005	0.30	0.11 to 0.56
<i>P</i> < 0.4	1.13	0.035	1.64	< 0.001	1.13	0.010	0.28	0.08 to 0.55
<i>P</i> < 0.5	1.11	0.081	1.64	< 0.001	1.12	0.018	0.24	0.04 to 0.50
Severe major depressive disorder^d								
<i>P</i> < 0.001	1.02	0.805	1.69	< 0.001	1.07	0.185	0.09	−0.08 to 0.28
<i>P</i> < 0.01	1.11	0.116	1.69	< 0.001	1.11	0.054	0.21	0.02 to 0.46
<i>P</i> < 0.05	1.22	0.002	1.69	< 0.001	1.14	0.013	0.37	0.14 to 0.72
<i>P</i> < 0.1	1.2	0.005	1.69	< 0.001	1.14	0.008	0.36	0.13 to 0.69
<i>P</i> < 0.2	1.17	0.016	1.69	< 0.001	1.13	0.017	0.33	0.10 to 0.67
<i>P</i> < 0.3	1.17	0.017	1.69	< 0.001	1.16	0.005	0.36	0.13 to 0.69
<i>P</i> < 0.4	1.17	0.016	1.69	< 0.001	1.14	0.009	0.34	0.11 to 0.70
<i>P</i> < 0.5	1.15	0.032	1.69	< 0.001	1.14	0.014	0.30	0.07 to 0.63

OR, odds ratio.
 a. Direct effects of the PRSs, childhood trauma and their interaction effects were estimated in four separate logistic regression models. The effects of the PRS (model 1) and childhood trauma (model 2) were estimated in models with age, gender and three principal components as covariates.
 b. The interaction effects were estimated in a model additionally including PRS and childhood trauma as covariates (model 3 and model 4). The relative excess risks due to interaction (RERI) represent tests for interaction as departure from additivity and were computed by $e^{BCT+PRS+PRSPCT} - e^{BCT} - e^{PRSPCT} - 1$.
 c. 1645 cases and 340 controls.
 d. 956 cases and 340 controls.

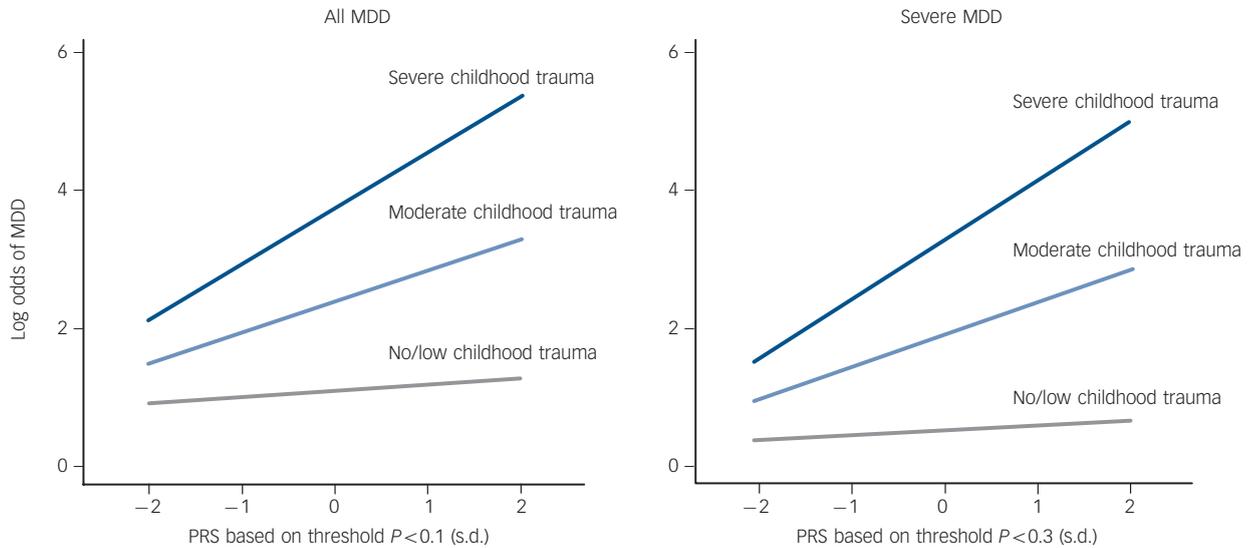


Fig. 1 Interaction between childhood trauma and polygenic risk score (PRS) on the risk for major depressive disorder (MDD).
 The interaction effects as departure of multiplicativity in predicting risk on all MDD and risk on severe MDD are visualised by displaying the direct effects of the PRS based on threshold *P* < 0.1 and *P* < 0.3 respectively for three childhood trauma levels, with childhood trauma scores of 0–1, 2–3 and 4–8 respectively.

Thus far, polygenic information has not been taken into account in research on G × E interaction in MDD, but there has been ongoing research for interaction with candidate genes. The motivation for research on G × E interaction in MDD is found in its contribution to understanding the complex aetiology of MDD,³³ and its possibility to select environmental conditions with increased genetic effects. Nevertheless, research on candidate genes has led to rather contradictory results: in research on the well-known serotonin transporter gene (5-HTTLPR) even meta-analyses differ in their conclusions,^{5,6,34} with concerns about publications bias.¹ However, because genetic effects on MDD are

polygenic in nature,^{11,12} we argued that G × E interaction should be tested with polygenic information.

The interaction effect thus found within our sample between polygenic risk scores and childhood trauma in MDD has two implications. The first is that polygenic risk scores have increased effects in the presence of childhood trauma (as illustrated in Fig. 1), which indicates that research on direct genetic effects potentially gains power by focusing on individuals exposed to childhood trauma. Therefore, if numbers would allow, it would be very useful to perform a GWAS within, for example, the collaborative PGC¹⁵ in individuals who experienced childhood

Table 3 Interaction between polygenic risk score (PRS) and four childhood trauma domains in predicting all major depressive disorder risk and direct effects of the four childhood trauma domains ($n = 1645$ cases, $n = 340$ controls)^a

PRS thresholds	Emotional neglect (EN)						Psychological abuse (PsA)						Physical abuse (PhA)						Sexual abuse (SA)					
	EN		PRS × EN		PsA		PRS × PsA		PhA		PRS × PhA		SA		PRS × SA									
	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P								
$P < 0.001$	2.57	<0.001	1.10	0.307	2.40	<0.001	1.03	0.809	2.90	<0.001	1.24	0.297	2.19	<0.001	1.12	0.503								
$P < 0.01$	2.57	<0.001	1.15	0.130	2.40	<0.001	1.16	0.216	2.90	<0.001	1.30	0.215	2.19	<0.001	1.05	0.785								
$P < 0.05$	2.57	<0.001	1.15	0.128	2.40	<0.001	1.37	0.006	2.90	<0.001	1.42	0.081	2.19	<0.001	1.08	0.653								
$P < 0.1$	2.57	<0.001	1.18	0.069	2.40	<0.001	1.36	0.007	2.90	<0.001	1.36	0.137	2.19	<0.001	1.11	0.508								
$P < 0.2$	2.57	<0.001	1.21	0.032	2.40	<0.001	1.29	0.025	2.90	<0.001	1.23	0.288	2.19	<0.001	1.10	0.561								
$P < 0.3$	2.57	<0.001	1.22	0.027	2.40	<0.001	1.36	0.007	2.90	<0.001	1.18	0.381	2.19	<0.001	1.12	0.460								
$P < 0.4$	2.57	<0.001	1.21	0.035	2.40	<0.001	1.32	0.016	2.90	<0.001	1.11	0.577	2.19	<0.001	1.14	0.393								
$P < 0.5$	2.57	<0.001	1.19	0.056	2.40	<0.001	1.31	0.018	2.90	<0.001	1.10	0.609	2.19	<0.001	1.16	0.356								

OR, odds ratio.

a. Four childhood trauma domains (ranging from 0 to 2) were tested for their direct effects and interaction effects with polygenic risk scores on all major depressive disorder risk, using logistic regression analyses adjusted for age, gender and three principal components. The main effects of the PRS are displayed in Table 2 (model 1).

trauma. Because interaction effects are symmetrical, we could, however, also have illustrated that childhood trauma has more impact in individuals with high polygenic risk scores. Thus, the second implication is that individuals with high polygenic risk scores are more vulnerable for the effects of childhood trauma, which has potential clinical relevance, for example in profiling of MDD, but also in possible future prevention programmes. When replicated in independent samples, the interaction effect found might add a modest but important piece to the complex puzzle of MDD's aetiology.

The direct effects of the polygenic risk scores and childhood trauma in predicting MDD risk in our sample are in line with previous findings. The proportion of variation in MDD explained by the polygenic risk scores ($R^2 \sim 0.5\%$) was in agreement with the findings of Demirkan *et al*¹¹ and the PGC.¹⁵ Although Nagelkerke's R^2 could have suffered from ascertainment bias because of the large proportion of participants with MDD in our sample, its estimates were of the same magnitude as Lee's estimates of R^2 , indicating that ascertainment did not largely affect our results.³⁰ The choice of the SNP P -value cut-off in the discovery sample tends to be arbitrary, which is why we presented results for eight different cut-offs in this study, and results were comparable for cut-offs larger than 0.05. In general, we anticipate that lower cut-offs are preferable over higher cut-offs when the discovery sample size increases and SNP effects can be found with more certainty. The impact of childhood trauma in predicting MDD risk in our sample is also in line with studies by MacMillan *et al*⁹ and De Graaf *et al*.⁷ Furthermore, evidence for interaction was found as departure from both multiplicativity and additivity, the latter of which has been argued to be more in line with biological interaction.^{29,35}

The impact of polygenic risk scores on MDD could have been studied in several environmental conditions, but we hypothesised that the presence of childhood trauma is a likely candidate. The presence of childhood trauma showed most consistent results in previous research on interaction with candidate genes,⁵ and it is a severe form of stress with a large and life-long impact, resulting in a large main effect on MDD prevalence.^{8,10} Furthermore, childhood trauma generally occurs before the onset of MDD (in our sample 84.7% of participants with MDD had their first episode after age 16), thereby largely excluding the potential source of bias from reciprocal causation, i.e. when MDD results in environmental stress.³⁶ In our study, childhood trauma was assessed with the CTI, which is a well-established instrument that has shown to predict onset of depressive and anxiety disorders^{7,20} as well as an enduring impact on structural and functional brain abnormalities.^{21,22} Our finding that childhood trauma increases the effects of polygenic risk scores on MDD fits with a recent review of Teicher & Samson, which indicates that patients with MDD and childhood trauma have more severe mood, neuro-vegetative and endogenous symptoms, and more comorbidities and psychotic features than patients with MDD without childhood trauma.¹⁰

The approach applied in this study, to test for $G \times E$ interaction with polygenic risk scores, has both advantages and disadvantages. This is the first study to apply this approach to MDD, but Meyers and colleagues have previously applied it to smoking behaviour. They observed interaction effects on smoking behaviour between polygenic risk scores for smoking and the number of traumatic events experienced as well as for polygenic risk scores and neighbourhood social cohesion (effective $n = 399$).³⁷ An advantage of the polygenic risk scores approach is that polygenic risk scores are based on genome-wide SNP data, but result in a one-dimensional summary measure, with corresponding requirements of significance ($P < 0.05$). Consequently, the sample

Table 4 Proportion of variation in risk on all major depressive disorder explained by childhood trauma, polygenic risk score (PRS) and their interaction effect

PRS thresholds	Nagelkerke's R^2 , ^a %			
	Childhood trauma	PRS	PRS × childhood trauma	Lee's R^2 , ^b PRS
$P < 0.001$	12.68	< 0.00	0.09	0.02
$P < 0.01$	12.68	0.30	0.24	0.41
$P < 0.05$	12.68	0.90	0.53	1.26
$P < 0.1$	12.68	0.66	0.60	0.94
$P < 0.2$	12.68	0.49	0.46	0.64
$P < 0.3$	12.68	0.40	0.59	0.54
$P < 0.4$	12.68	0.40	0.49	0.56
$P < 0.5$	12.68	0.26	0.42	0.39

a. To estimate proportions of variation in risk on all major depressive disorder explained by childhood trauma, PRS and their interaction (PRS × childhood trauma), Nagelkerke's R^2 were compared: between the model with only covariates and the model additionally including childhood trauma (R^2 of childhood trauma); between the model with covariates and childhood trauma and the model additionally including PRS (R^2 of PRS); and between the model with covariates, childhood trauma and PRS and the model additionally including the interaction PRS × childhood trauma (R^2 of PRS × childhood trauma). Age, gender and three principal components were included as covariates.

b. Lee's proposed estimate of R^2 was computed³⁰ using a Dutch lifetime prevalence of major depressive disorder of 18.7%.³¹

size of the target sample can be much smaller than in GWAS testing SNPs independently. A disadvantage is, however, that particular aspects of the multidimensional polygenic information are lost, which could lead to biased results, for example when certain SNPs show increased effects on MDD in the presence of childhood trauma, whereas other SNPs show decreased effects on MDD in the presence of childhood trauma. If this hypothetical situation would occur, both interaction effects would be levelled out in tests with the one-dimensional polygenic risk scores summary measure. Nevertheless, at the present time sample sizes are insufficient to examine the impact of many SNPs independently in $G \times E$ studies and, therefore, studying polygenic risk scores seems an elegant approach for testing the general hypothesis of the existence of $G \times E$ interaction.

Strengths and limitations

Our study has several strengths. First, it was based on DSM-IV diagnoses of MDD, which ensures we studied participants with clinically relevant MDD. Second, controls were carefully screened for any lifetime psychiatric diagnosis. Third, childhood trauma was assessed in a face-to-face interview by specially trained clinical staff. Fourth, polygenic risk scores were based on a large and independent discovery sample, which adds to the accuracy of the polygenic risk scores. However, there are also some limitations, including potential recall bias of childhood trauma influenced by the mood of participants with MDD. The number of controls in our sample was rather limited, but we carefully checked for ascertainment bias and found none. Even though controls were carefully screened for MDD, they could potentially develop MDD later in life, especially because MDD has a high prevalence of approximately 15–20%.³¹

Clinical implications

We show that the effect of polygenic risk scores on MDD is increased in the presence of childhood trauma in our sample. Our finding implicates that power in research on direct genetic effects is larger in the presence of childhood trauma, but it also implicates that individuals with high polygenic risk scores form a potential group for MDD prevention, because of their increased vulnerability to the depressogenic effects of childhood trauma. Future research should be conducted to replicate our finding, especially in the light of the inconclusive findings in research on interaction in MDD thus far. In addition, future research could

also be designed to test interaction with polygenic information applying different techniques. A possible technique to apply could be genome-wide complex trait analyses (GCTA) to test for interaction with the genetic relationship matrix.³⁸ The present study was underpowered to conduct such analyses,³⁹ but future efforts combining data from several independent GWAS cohorts could potentially reach power to test for interaction with GCTA. Further research is required, but our results suggest that $G \times E$ interaction could play a considerable role in the polygenic effects on MDD.

Funding

The Netherlands Study of Depression and Anxiety (NESDA) was funded by the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CMSB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, ENGAGE (HEALTH-F4-2007-201413). Genotyping was funded by the US National Institute of Mental Health (NIMH) (RC2MH089951) as part of the American Recovery and Reinvestment Act of 2009. The Psychiatric Genomics Consortium was funded by NIMH grants MH085520 and MH080403. A.A. and W.J.P. were supported by CSMB/Neuroscience Campus Amsterdam. Statistical analyses were partly carried out on the Genetic Cluster Computer (www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003), the Dutch Brain Foundation, and the Department of Psychology and Education of the VU University Amsterdam.

Wouter J. Peyrot, MD, **Yuri Milaneschi**, PhD, Department of Psychiatry, Neuroscience Campus Amsterdam and EMGO Institute for Health and Care Research, VU University Medical Center & GGZ inGeest, Amsterdam, The Netherlands; **Abdel Abdellaoui**, MSc, Department of Biological Psychology, VU University Amsterdam, and Neuroscience Campus Amsterdam, Amsterdam, The Netherlands; **Patrick F. Sullivan**, PhD, Departments of Genetics and Psychiatry, University of North Carolina at Chapel Hill, North Carolina, USA; **Jouke J. Hottenga**, PhD, **Dorret I. Boomsma**, PhD, Department of Biological Psychology, VU University Amsterdam; Neuroscience Campus Amsterdam and EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands; **Brenda W. J. H. Penninx**, PhD, Department of Psychiatry, Neuroscience Campus Amsterdam and EMGO Institute for Health and Care Research, VU University Medical Center & GGZ inGeest, Amsterdam, The Netherlands

Correspondence: Wouter J. Peyrot, MD, Department of Psychiatry, VU University Medical Center & GGZ inGeest, AJ Ernststraat 1187, 1081 HL Amsterdam, The Netherlands. Email: w.peyrot@ggzingeest.nl

First received 7 Dec 2013, final revision 6 Mar 2014, accepted 14 Mar 2014

References

- Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry* 2011; **168**: 1041–9.
- Klengel T, Binder EB. Gene-environment interactions in major depressive disorder. *Can J Psychiatry* 2013; **58**: 76–83.
- Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2003; **301**: 386–9.

- 4 Fergusson DM, Horwood LJ, Miller AL, Kennedy MA. Life stress, 5-HTTLPR and mental disorder: findings from a 30-year longitudinal study. *Br J Psychiatry* 2011; **198**: 129–35.
- 5 Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis: evidence of genetic moderation. *Arch Gen Psychiatry* 2011; **68**: 444–54.
- 6 Risch N, Herrell R, Lehner T, Liang KY, Eaves L, Hoh J, et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA* 2009; **301**: 2462–71.
- 7 De Graaf R, Bijl RV, Smit F, Vollebergh WA, Spijker J. Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: findings from the Netherlands Mental Health Survey and Incidence Study. *Am J Psychiatry* 2002; **159**: 620–9.
- 8 Hovens JGFM, Wiersma JE, Giltay EJ, van Oppen P, Spinhoven P, Penninx BWJH, et al. Childhood life events and childhood trauma in adult patients with depressive, anxiety and comorbid disorders vs. controls. *Acta Psychiatr Scand* 2010; **122**: 66–74.
- 9 MacMillan HL, Fleming JE, Streiner DL, Lin E, Boyle MH, Jamieson E, et al. Childhood abuse and lifetime psychopathology in a community sample. *Am J Psychiatry* 2001; **158**: 1878–83.
- 10 Teicher MH, Samson JA. Childhood maltreatment and psychopathology: a case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. *Am J Psychiatry* 2013; **170**: 1114–33.
- 11 Demirkan A, Penninx BW, Hek K, Wray NR, Amin N, Aulchenko YS, et al. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol Psychiatry* 2011; **16**: 773–83.
- 12 Lubke GH, Hottenga JJ, Walters R, Laurin C, de Geus EJ, Willemsen G, et al. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biol Psychiatry* 2012; **72**: 707–9.
- 13 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.
- 14 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger Jr, et al. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–9.
- 15 Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.
- 16 Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 2011; **43**: 519–25.
- 17 Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–94.
- 18 Penninx BW, Beekman AT, Smit JH, Zitman FG, Nolen WA, Spinhoven P, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* 2008; **17**: 121–40.
- 19 Penninx BW, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, et al. Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord* 2011; **133**: 76–85.
- 20 Peyrot WJ, Middeldorp CM, Jansen R, Smit JH, de Geus EJ, Hottenga JJ, et al. Strong effects of environmental factors on prevalence and course of major depressive disorder are not moderated by 5-HTTLPR polymorphisms in a large Dutch sample. *J Affect Disord* 2013; **146**: 91–9.
- 21 van Harmelen AL, van Tol MJ, van der Wee NJ, Veltman DJ, Aleman A, Spinhoven P, et al. Reduced medial prefrontal cortex volume in adults reporting childhood emotional maltreatment. *Biol Psychiatry* 2010; **68**: 832–8.
- 22 van Harmelen AL, van Tol MJ, Demenescu LR, van der Wee NJ, Veltman DJ, Aleman A, et al. Enhanced amygdala reactivity to emotional faces in adults reporting childhood emotional maltreatment. *Soc Cogn Affect Neurosci* 2013; **8**: 362–9.
- 23 Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl* 2003; **27**: 169–90.
- 24 Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D, et al. Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur J Hum Genet* 2008; **16**: 335–42.
- 25 Abdellaoui A, Hottenga JJ, Knijff PD, Nivard MG, Xiao X, Scheet P, et al. Population structure, migration, and diversifying selection in the Netherlands. *Eur J Hum Genet* 2013; **21**: 1277–85.
- 26 Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* 2009; **14**: 359–75.
- 27 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.
- 28 Purcell S. Variance components models for gene-environment interaction in twin analysis. *Twin Res* 2002; **5**: 554–71.
- 29 Knol MJ, van der Tweel I, Grobbee DE, Numans ME, Geerlings MI. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *Int J Epidemiol* 2007; **36**: 1111–8.
- 30 Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol* 2012; **36**: 214–24.
- 31 De Graaf R, Ten Have M, Van Gool C, Van Dorsselaer S. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Soc Psychiatry Psychiatr Epidemiol* 2012; **47**: 203–13.
- 32 R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2010.
- 33 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747–53.
- 34 Munafò MR, Durrant C, Lewis G, Flint J. Gene × environment interactions at the serotonin transporter locus. *Biol Psychiatry* 2009; **65**: 211–9.
- 35 Uher R, Caspi A, Houts R, Sugden K, Williams B, Poulton R, et al. Serotonin transporter gene moderates childhood maltreatment's effects on persistent but not single-episode depression: replications and implications for resolving inconsistent results. *J Affect Disord* 2011; **135**: 56–65.
- 36 Middeldorp CM, Cath DC, Beem AL, Willemsen G, Boomsma DI. Life events, anxious depression and personality: a prospective and genetic study. *Psychol Med* 2008; **38**: 1557–65.
- 37 Meyers JL, Cerda M, Galea S, Keyes KM, Aiello AE, Uddin M, et al. Interaction between polygenic risk for cigarette use and environmental exposures in the Detroit neighborhood health study. *Transl Psychiatry* 2013; **3**: e290.
- 38 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.
- 39 Visscher PM, Hemani G, Vinkhuyzen AAE, Chen GB, Lee SH, Wray NR, et al. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genetics* 2014; **10**: e1004269.