

Exploring polygenic-environment and residual-environment interactions for depressive symptoms within the UK Biobank

Alexandra C. Gillett¹, Bradley S. Jermy^{1,2}, S. Hong Lee³, Oliver Pain^{1,2}, David M. Howard^{1,4}, Saskia P. Hagenaars¹, Ken B. Hanscombe¹, Jonathan R. I. Coleman^{1,2} and Cathryn M. Lewis^{1,2,5}.

¹Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

²NIHR Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, SE5 8AF, UK

³Australian Centre for Precision Health, UniSA Allied Health and Human Performance, University of South Australia, Adelaide, SA 5000, Australia

⁴Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK

⁵Department of Medical and Molecular Genetics, Faculty of Life Sciences and Medicine, King's College London

Corresponding author:

Dr. Alexandra Gillett
Social, Genetic and Developmental Psychiatry Centre
Institute of Psychiatry, Psychology and Neuroscience - PO80
De Crespigny Park, Denmark Hill, London,
United Kingdom, SE5 8AF
Phone: 02078485299
Email: alexandra.a.gillett@kcl.ac.uk

Keywords: genotype-environment interaction, residual-environment interaction, depressive symptoms, multivariate reaction norm model

Abstract

Substantial advances have been made in identifying genetic contributions to depression, but little is known about how the effect of genes can be modulated by the environment, creating a gene-environment interaction. Using multivariate reaction norm models (MRNMs) within the UK Biobank (N=61294-91644), we investigate whether the polygenic and residual variation of depressive symptoms are modulated by 25 a-priori selected covariate traits: 12 environmental variables, 5 biomarkers and polygenic risk scores for 8 mental health disorders. MRNMs provide unbiased polygenic-covariate interaction estimates for a quantitative trait by controlling for outcome-covariate correlations and residual-covariate interactions. Of the 25 selected covariates, 11 significantly modulate depressive symptoms, but no single interaction explains a large proportion of phenotypic variation. Results are dominated by residual-covariate interactions, suggesting that covariate traits (including neuroticism, childhood trauma and BMI) typically interact with unmodelled variables, rather than a genome-wide polygenic score, to influence depressive symptoms. Only average sleep duration has a polygenic-covariate interaction explaining a demonstrably non-zero proportion of the variability in depressive symptoms. This effect is small, accounting for only 1.22% (95% CI [0.54,1.89]) of variation. The presence of an interaction highlights a specific focus for intervention, but the negative results here indicate a limited contribution from polygenic-environment interactions.

Introduction

Major depressive disorder (MDD) is a common and debilitating mental disorder which is the second leading cause of years lived with disability worldwide (Vos, et al., 2020). It has a lifetime prevalence of 17.8% in global populations (Vos, et al., 2016). The core symptoms of depression are persistent low mood and anhedonia, with other diagnostic signs and symptoms including changes in cognition, appetite, or sleep, and feelings of fatigue and worthlessness. The heritability of MDD is lower than many other psychiatric disorders, estimated at between 30-40%, with higher values for severe cases (Kendall, et al., 2021). This lower heritability suggests that a substantial proportion of liability to depression is due to environmental risk factors.

Genome-wide association studies (GWAS) have made progress in identifying variants associated with MDD, with 178 loci now identified and SNP-heritability estimates ranging between 5.5% and 11.2%, depending on the depression definition used (Levey, et al., 2021). The difference between pedigree and SNP-heritability estimates may indicate a role for additional sources of genetic related variation, such as gene-environment (G-E) interactions. Identifying G-E interactions would provide insight into the biological mechanisms of depression, improve the accuracy of heritability estimates and path the way to individualised preventative healthcare (Hunter, 2005). Genetic studies have used a wide range of definitions of MDD, from diagnosis in clinical studies, to self-report of a diagnosis with depression, to reported presence of depressive symptoms. These criteria show a strong common genetic overlap, with pairwise genetic correlations of at least 0.7 for most MDD definitions (Jermy, et al., 2020; Levey, et al., 2021). We have previously shown that a continuous measure of depression, based on factor analysis of questionnaire responses in the general population effectively captures the polygenetic component of depression (Jermy, et al., 2020).

As noted above, a substantial component of the liability to depression arises from nongenetic factors. Stressful life events and exposure to trauma provide the strongest risks, with smoking, obesity, BMI and exercise also associated with depression (Luppino, et al., 2010; Gianfredi, et al., 2020; Coleman, et al., 2020). The correlation or interaction between these risk factors and the genetic predisposition to depression has been largely un-investigated. Performing studies to disentangle the genetic and environmental contributions to complex traits such as depression is challenging. The environmental variables potentially have a genetic component, and these traits may also be genetically correlated with depression (Wray, et al., 2018). Causation might be multi-directional, with depression risk increased as a consequence of a risk factor, or a risk factor being observed because a healthy lifestyle is more challenging to maintain during a depressive episode. Further, depression may influence the reporting of risk factor status, for example, retrospective reporting of trauma differs from prospective reporting (Baldwin, Reuben, Newbury, & Danese, 2019). These complexities make testing for G-E interactions and correlations challenging, as highlighted by studies investigating an interaction between polygenic risk scores (PRSs) for depression and reported childhood trauma. Early investigations identified an interaction (Peyrot, et al., 2014; Mullins, et al., 2016), but a larger study in the Psychiatric Genomics Consortium (PGC) found no evidence for departure from additive contribution to risk of depression (Peyrot, et al., 2018).

In this paper we model a continuous measure of depressive symptoms and explore genome-wide genotype-covariate (G-C) and residual-covariate (R-C) interactions for 25 covariate traits, including environmental risk factors. A significant G-C interaction means that the additive genetic component for symptoms of depression (G), which has been estimated internal to the data, varies with respect to a covariate trait (C) (Xuan, et al., 2020). This can be thought of as a polygenic-covariate interaction (Dahl, et al., 2020). A significant R-C interaction means that the variation observed in symptoms of depression is modulated by the covariate trait, but in a manner not specified by the model; hence it is a residual interaction. We analyse 17 measured traits in UK Biobank (UKB) including BMI and related body composition traits, exercise measures, smoking, neuroticism, sleep duration, childhood trauma, Townsend deprivation index (TDI) and biomarkers. We additionally analyse PRSs for eight psychiatric disorders. The interactions are modelled in a reaction norm analysis using mtg2 software (Lee & van der Werf, 2016) which tests whether individual differences in the genetic and residual effects are modulated by another risk factor. The multivariate reaction norm model (MRNM) uses covariance functions to model interactions between high-dimensional sets of genetic variants and environmental covariates whilst controlling for trait correlations. It is useful when it is not feasible to investigate interactions variant by variant due to dimensionality (Jarquin, et al., 2014), and has higher power compared to single variant interaction tests for polygenic traits (Dahl, et al., 2020). This statistical framework allows us to robustly investigate the role of these factors in modulating the polygenic, and residual, effects on depression symptoms.

Methods

UK Biobank (UKB)

Analysis was performed using the UKB, a health study of over 500,000 UK participants who were recruited in mid-life (40-69 years old) between 2006 and 2010 (Sudlow, et al., 2015). Detailed information on health and lifestyle are available from self-report at baseline, when biological samples for genetic analysis and biomarker testing were also taken. A follow up Mental Health Questionnaire (MHQ), completed online by 157,339 participants in 2016, collected information on a wide range of lifetime psychiatric diagnoses and current depression symptoms (Davis, et al., 2020).

Outcome of interest

An outcome trait measuring depression symptoms was derived from the MHQ assessment of depressive symptoms over the last two weeks, which are drawn from the PHQ-9 and correspond to the Diagnostic and Statistical Manual of Mental Disorders criteria for MDD. This trait, which we call depSympt, was constructed via a hierarchical model by Jermy et al. (2020) using MHQ depression-related symptom data. It summarises symptoms related to mood, anxiety, subjective well-being, psychomotor cognitive factors and neuro-vegetative factors. A summary of depSympt and its construction can be found in the Supplementary Materials (SM) section 1.1, with full information given in Jermy et al. (2020).

depSympt is strongly associated with lifetime MDD status, defined using the Composite International Diagnostic Interview Short Form, where it explains 11% of the variation in liability to MDD (Jermy, et al., 2020). On average, prevalent MDD cases have higher depSympt values compared to controls, showing that ever having had depression is associated with increased current depressive symptoms compared to never having had depression (Supplementary Figure 2). Permutation tests (used as depSympt is non-normal) showed highly significant differences between both mean and median depSympt values in MDD cases and controls ($p < 1E - 315$).

The covariate traits

25 covariate traits were selected, based on previous associations with MDD phenotypes and availability in UKB. These comprise 17 environmental and biomarker covariate traits. Body composition was represented by body mass index (BMI), waist-to-hip ratio, and waist circumference. Exercise variables used the metabolic equivalent task (MET) scores based on the International Physical Activity Questionnaire, which assesses the frequency, intensity and duration of exercise in three categories: walking, moderate exercise and vigorous exercise. Four variables of summed MET minutes per week were analysed: all activities (MET total), and the separate categories of walking (MET walk), moderate exercise (MET mod), and vigorous exercise (MET vig). Other covariate traits from the baseline assessment were Townsend deprivation index (TDI), average sleep duration (sleep), neuroticism score, and pack years of smoking (smoking). Five biomarkers of LDL cholesterol, HDL cholesterol, triglycerides, C-reactive protein (CRP) and vitamin D were analysed. All biomarkers except for LDL were log transformed (see SM section 1.3). From the MHQ, a continuous variable

summarising reported childhood trauma was analysed (Pitharouli, et al., 2021). All covariates were from the baseline UKB assessment, except reported childhood trauma, which was collected in the MHQ.

PRSs for the following eight mental health traits were also selected as covariates: major depressive disorder (MDD) (Wray, et al., 2018), bipolar disorder (BIP) (Sklar, et al., 2011), schizophrenia (Ripke, et al., 2014), anxiety (Otowa, et al., 2016), attention deficit hyperactivity disorder (ADHD) (Demontis, et al., 2019), autism spectrum disorder (ASD) (Anney, et al., 2017), obsessive compulsive disorder (OCD) (Arnold, et al., 2018) and anorexia nervosa (Watson, et al., 2019). Construction of the scores is described in SM section 1.2, and details of the GWAS summary statistics are given in Supplementary Table 5.

For the 17 environmental traits, the G-C and the R-C interactions are gene-environment and residual-environment interactions respectively. For the eight PRSs, the G-C interactions represent a narrow form of gene-gene interaction which captures variation in depressive symptoms attributable to an interaction between the PRS and the residual additive genetic component of depression symptoms after removing the fixed effect of the PRS.

Genetic data

Autosomal genotype data underwent a centralised quality control procedure described by Bycroft et al. (2018) prior to imputation. We then selected HapMap3 single nucleotide polymorphisms (SNPs) from the imputed UKB genetic data (Ni, et al., 2019; Xuan, et al., 2020), and further removed variants with a minor allele frequency < 0.01, an information score (used to index the quality of genotype imputation) < 0.7 and completeness < 95% (Coleman, et al., 2016).

Quality control for participants followed procedures detailed by (Coleman, et al., 2020). Briefly, analysis was limited to unrelated individuals of European ancestry who had completed the online MHQ, had a call rate of >98% for genotyped SNPs and for whom genetic sex matched self-reported sex. Additionally, individuals were removed for unusual levels of missingness or heterozygosity where recommended by the UKB core analysis team, or if they had withdrawn consent for analysis. After quality control, 126,522 participants were retained. This reduced to 119,692 after omitting individuals missing the outcome trait, depSympt (Jermy, et al., 2020). From the 1,118,287 SNPs retained, genetic relationship matrices (GRMs) were created using Plink version 1.9 for use in the interaction models (Yang, Lee, Goddard, & Visscher, 2011; Chang, et al., 2015).

Statistical Analysis

Interaction analysis was performed using a mixed effects model called the multivariate reaction norm model (MRNM) (Ni, et al., 2019), which includes two interaction types (polygenic and residual) as random effects, and adjusts for genetic and residual correlations between the outcome and covariate traits. In the sections that follow we provide a broad overview of the MRNM and detail its application in this work, with Figure 1 summarising our approach within the UKB.

Model overview

A reaction norm (RN) is a genotype-specific function describing the relationship between an outcome and a covariate trait. Interactions are indicated by non-parallel RNs which produce a relationship between the variability of outcome and the covariate within a population. The MRNM looks for evidence of interactions by estimating the trend in outcome variability across the covariate trait, and decomposing this into polygenic and residual sources using estimated genetic similarities from genome-wide SNP data within a random effects model (Schaeffer, 2004; Jarquin, et al., 2014; Ni, et al., 2019). Genetic similarity here is defined using the GRM.

The MRNM has two modelling stages. Firstly, in the fixed effects model, linear regression is used to estimate the expected value of a trait using a set of variables (fixed effects) selected for inclusion (e.g. genotype batch to adjust for possible confounding). This is done for both the outcome trait (here, depSympt) and the covariate traits (see SM section 1.3 for details). Secondly, using a multivariate random effects model, the standardised residual variation in outcome (Y) and the covariate trait (C), not explained by their respective fixed effects models, is partitioned into genetic and residual random components. For the outcome trait these genetic and residual random effects can be a function of C , allowing heterogeneity of the polygenic and residual variance components for Y across C , thereby incorporating genotype-covariate (G-C) and residual-covariate (R-C) interactions into the model. This random effects model can be written as:

$$\begin{bmatrix} Y_i | C_i = c_i \\ C_i \end{bmatrix} = \begin{bmatrix} \alpha_{0i} + \alpha_{1i}c_i \\ \beta_{0i} \end{bmatrix} + \begin{bmatrix} \tau_{0i} + \tau_{1i}c_i \\ \varepsilon_{0i} \end{bmatrix} \quad (1)$$

where, for an individual i : 1. α_{0i} (β_{0i}) is a random effect defining the random polygenic intercept for Y_i (C_i), 2. τ_{0i} (ε_{0i}) is a random effect defining the random residual intercept for Y_i (C_i), 3. α_{1i} is a random effect capturing the interaction of the polygenic component for Y_i with C_i (the G-C interaction), and 4. τ_{1i} is a random effect capturing the trend in the phenotypic variability of Y_i across C_i that is not explained by the measured genetic variables (the R-C interaction, which can be thought of as covariate-specific noise). These random effects are random variables, characterised by a population multivariate normal distribution with mean zero, and a covariance matrix requiring estimation. The MRNM is therefore parameterised by estimating the covariance matrix between the random effects, which represent sources of (co)variation for Y and C within the population. Genetic (residual) correlation between the two traits is included in the model by estimating the covariance parameters between the genetic (residual) random effects for Y and for C . This is an advantage of the MRNM because not accounting for trait correlations can lead to an inflation in the strength of the interactions (Ni, et al., 2019).

Using the MRNM in Equation (1), the marginalised variance for Y_i is:

$$Var[Y_i] = \sigma_{\alpha_0}^2 + \sigma_{\alpha_1}^2 + \sigma_{\tau_0}^2 + \sigma_{\tau_1}^2 = 1 \quad (2)$$

for all individuals in the population. $\sigma_{\alpha_0}^2$, $\sigma_{\alpha_1}^2$, $\sigma_{\tau_0}^2$ and $\sigma_{\tau_1}^2$ therefore measure the importance of the effect they capture (homogeneous polygenic component, G-C interaction,

homogeneous residual component and R-C interaction) in explaining the variability of outcome in the population.

Phenotype adjustment

Interactions are assessed for each covariate trait in turn. In the first modelling stage of the MRNM, we adjust depSympt and the covariate trait for fixed effects, using a linear model. Fixed effects included demographic variables (age, sex, and assessment centre) and population structure using the first 15 principal components. We also adjust for stressful and traumatic events in adulthood, due to their likely impact on depressive symptoms. SM section 1.3 details the fixed effects variables used for each depSympt-covariate trait combination.

Except for principal components, continuous fixed effects variables were allowed to have a non-linear relationship with depSympt and the covariate traits by using fractional polynomials (FPs) (Royston & Altman, 1994). We used the mfp package (Benner & Ambler, 2015) within a generalised linear model `stats::glm` (R Core Team, 2020).

After fixed effects adjustment, all traits are standardised to allow comparison of their relative importance in explaining the variability in outcome across interaction models. For depSympt, we can recover the size of an interaction effect on the original, rather than residualised, scale and use the MRNMs to provide an estimate of the proportion of variability in depSympt explained by G-C and R-C interactions.

Covariate traits were analysed in five groups, each with a different sample size based on the missingness of covariate traits (see Figure 1). This addressed the trade-off between maximising sample size for each depSympt-covariate trait interaction analysis and the computational burden of constructing GRMs for each analysis. The fixed effects models are built using all available data within an analysis group, with sample sizes ranging from 61,294 to 91,644.

Identifying interactions

For each (fixed effects adjusted and standardised) covariate trait, we use MRNMs (Ni, et al., 2019) to evaluate evidence of interactions explaining a non-zero proportion of the variability in depSympt. Following the approach of Xuan, et al. (2020), we run a full MRNM, with both G-C and R-C interactions included as random effects, and a null MRNM with no interactions. A likelihood ratio test (LRT) is then used to compare these models, with a significant LRT Bonferroni-corrected p-value ($p < 0.05/25 = 0.002$) providing evidence that interactions explain a non-zero proportion of outcome variance.

This approach assesses the evidence for an overall interaction effect. Simulation studies have shown there is low power to disentangle G-C and R-C interactions in nested MRNM comparisons (Xuan, et al., 2020), with biased G-C interaction estimates to be expected if unmodelled R-C interactions are present (Dahl, et al., 2020). In contrast, the full MRNM produced unbiased estimates of G-C and R-C variance components (Ni, et al., 2019; Xuan, et al., 2020). For significant covariate traits, we can therefore use variance component

estimates from the full model, with 95% confidence intervals, to identify which interaction type, polygenic and/ or residual, explain the variability in depSympt.

MRNMs are computationally demanding. Therefore, to perform interaction analysis for a given covariate trait at biobank scale we randomly divided the available UKB participants into three subgroups. MRNMs were run, using the mtg2 package (Lee & van der Werf, 2016), and compared within each subgroup. Results were then meta-analysed using Fisher's method (Evangelou & Ioannidis, 2013), as described in SM section 1.5. Fixed effects adjustments and post-modelling analysis, including creating graphics, was performed using R version 4.0.4. A Bonferroni correction was applied to adjust for multiple testing (giving a significance threshold of $p = 0.05/25 = 0.002$). A sensitivity analysis was performed by re-fitting the MRNMs using a rank-based inverse normal transformation (RINT) of depSympt. Applying this transformation can control the type I error rate when the assumption of normality is violated (Ni, et al., 2019) and loss of signal indicates spurious interaction effects in the untransformed model (Xuan, et al., 2020). Simulations have shown that parameter estimates from the full model, without applying the RINT, remain unbiased when the normality assumption is violated (Xuan, et al., 2020). The MRNM without transforming depSympt was therefore used for variance component estimation.

Results

Prior to interaction analysis, linear regression models for the outcome trait, depSympt, with each covariate trait in turn were run (Supplementary Table 8). With the exception of LDL and the PRS for OCD, all covariate traits considered have a statistically significant main effect ($p < 0.05/25 = 0.002$), providing evidence that expected current symptoms of depression vary with these covariate traits. The effect sizes vary widely, ranging from 0.014% of the variability in depSympt explained by the ASD PRS to 21.55% of the variability in depSympt explained by neuroticism score. When fractional polynomials (FPs) are considered, which allow covariate traits to have a non-linear relationship with the expected value of depSympt, the effect of LDL on expected depSympt also becomes statistically significant, although the proportion of variability in depSympt explained is low (0.04%). The proportion of phenotypic variability in depSympt explained by each covariate trait does not greatly change between the main effects and the FP linear models, with average sleep duration having the largest absolute increase from 0.87% in the main effects model, to 2.24% in the FP model (Supplementary Tables 9 and 10).

Distribution plots of the outcome variable summarising depressive symptoms, depSympt, and each covariate trait (before and after fixed effects adjustment) can be found in Supplementary Figures 3 - 27, with Supplementary Table 7 presenting the distribution characteristics for the unadjusted traits. Prior to fixed effects adjustment, the distribution of standardised depSympt is not normally distributed, with evidence of some positive skew (skewness = 0.47). Some deviations from normality are still present after fixed effects adjustment, with a median skewness of 0.34, however, we note that deviations from normality for depSympt after fixed effects adjustment does not mean that the normality assumption of the MRNM is violated since this applies to the distribution of outcome conditional on the fixed *and* the random effects models. Some covariate traits, for example MET total, are highly skewed even after fixed effects adjustment, however, our primary focus is on the variance components for the outcome trait, for which we performed a sensitivity analysis, re-fitting the MRNMs using the RINT for depSympt to control the type I error rate if the assumption of normality is violated.

For each of the 25 covariate traits considered, MRNMs with and without interactions were run for depSympt, and evidence of G-C and/ or R-C interactions was assessed using LRTs. 11 of the 25 covariate traits had p-values below the Bonferroni-corrected significance level, which are presented in Table 1. These 11 traits remained significant in the RINT-based sensitivity analysis (Supplementary Table 13), used to check significant results were not due to normality violations, providing evidence for an interaction effect between depSympt and the following variables: neuroticism, childhood trauma, average sleep duration, BMI, waist circumference, smoking, waist-to-hip ratio, TDI, summed MET minutes for all activities, for walking and for moderate activities. These results show that, even after adjusting for fixed effects, individual-level differences in these covariate traits contribute to variation in depSympt.

To assess if these covariate traits modulate the polygenic component and/ or the residual component of depression symptoms (the G-C and R-C interaction respectively), we use a measure of interaction strength: the proportion of variability in depSympt explained by an

interaction effect. This measure is plotted as a %, with 95% CIs, for: (a) the G-C interaction and (b) the R-C interaction in Figure 2. When a 95% CI includes zero, we cannot be certain that the interaction term explains any of the variability in depSympt.

Nine of the eleven significant covariate traits had 95% CIs for the percentage of variability in depSympt attributable to R-C interactions that excluded zero, with point estimates ranging from 1.03% (95% CI: 0.34-1.73) for waist to hip ratio, to 2.98% (95% CI: 2.18-3.77) for childhood trauma. These results show that a small, but significant, proportion of the residual variability in depSympt is modulated by the following covariate traits: neuroticism, childhood trauma, average sleep duration, BMI, waist circumference, smoking, waist-to-hip ratio, summed MET minutes per week walking and TDI.

For the G-C interactions, only average sleep duration had a 95% CI that excluded zero (Figure 2). This polygenic-sleep interaction is estimated to explain 1.22% of the variability in depSympt (95% CI: 0.54-1.89). Figure 3 plots the relationship between the variance components (polygenic, residual and total) for depSympt and fixed effects adjusted average sleep duration, as estimated by the full MRNM. It shows a U-shaped relationship between the polygenic variance component and sleep, suggesting a larger polygenic contribution to depression symptoms for individuals getting far more or less sleep than expected compared to those in central sleep duration percentiles (see Supplementary Figure 46).

For the 11 covariates with significant interactions, R-C and G-C interactions account for only a small proportion of the variation in depSympt, and significant results are primarily driven by R-C interactions, not G-C interactions. These results imply that covariate traits exert more influence on depressive symptoms through the modulation of residual (unmodelled) effects compared to polygenic effects.

We note the presence of negative variance component estimates, which are possible within random effects models including the MRNM (Ni, et al., 2019; Xuan, et al., 2020). These values arise as the algorithm estimates this parameter freely, without any constraint that the variance must be positive. All negative variance component estimates have 95% CIs that cross 0, indicating that they are under-estimates of a null interaction effect.

Whilst interactions involving the classic depression environmental risk factors were significant – including with childhood trauma, BMI, exercise measures and smoking – none of the PRSs for psychiatric disorders nor the biomarkers showed significant results for the G-C or the R-C interaction. The MDD PRS was close to the significance threshold ($p = 2.63E - 03 > 0.002$) and was below the threshold for the RINT-based analysis ($p = 7.72E - 04 < 0.002$), but the overall impact of both the G-C and R-C interactions, as measured by the proportion of variability in depSympt they explain, have 95% CIs overlapping with zero, in line with the null finding.

For two of the eleven covariate traits, MET total and MET moderate, where the interaction model provided a better fit to the data compared to the null model, the 95% confidence intervals for the proportion of variance explained overlap with 0 for both the R-C and the G-C interactions. These results imply that while there is evidence for a trend in the variability

of depSympt across these covariate traits, we are unable to confidently disentangle the source of this interaction into genetic or residual components.

Discussion

G-E interactions in depression may give insights into its etiology, highlighting both the biological mechanism and identifying environmental risk factors whose role in depression is moderated by genetics. The low heritability for depression compared to many other psychiatric disorders, and the prominent role of environmental risk factors through trauma and stress, make depression a natural target for exploring G-E interactions. Previous G-E interaction studies in depression using genome-wide data, performing SNP-by-environment tests, have identified few significant and replicated results either in European ancestries (Arnau-Soler, et al., 2019) or in Hispanic, African American and Hans Chinese populations (Dunn, et al., 2016; Peterson, et al., 2018). Increasing sample sizes may yield more significant SNP-by-environment interactions, but the highly polygenic nature of depression makes searching for polygenic-environment interactions appealing. Investigations of G-E interactions using depression PRSs have demonstrated null or conflicting results (Kendall, et al., 2021), and these PRS-environment interaction analyses have not modelled a residual trend in the variability of outcome, which can bias G-E interaction estimates (Dahl, et al., 2020). Additionally, the SNP effects used to construct a PRS are estimated assuming an additive genetic model. PRSs are therefore currently designed with the expectation that they do not vary across an environmental gradient. To assess whether the polygenic component for a trait is modified by the environment, a model which can create a polygenic score allowing for this possibility should be employed.

Furthermore, for a complex trait like depression, we would expect the genetic space to map to the phenotypic space through the environment, making the outcome a complex interplay between genes and environment (Assary, Vincent, Machlitt-Northen, Keers, & Pluess, 2020). Since environmental traits are often complex traits themselves, part of this interplay may include a genetic correlation between environment and depression, i.e., BMI, smoking and exercise measures have a positive genetic correlation with depression (Wray, et al., 2018). Not accounting for G-E correlation whilst investigating G-E interaction can lead to biased variance component estimates (Purcell, 2002), and have been shown to inflate the significance of interaction results (Ni, et al., 2019). Interaction analyses using PRSs do adjust for correlation through a main effect, however, they do not allow for residual correlations between traits nor offer the opportunity to investigate G-E interactions in the *presence* of G-E correlation. In this paper we have used MRNMs (Ni, et al., 2019) within the UKB to identify genome-wide genotype-covariate (G-C) and residual-covariate (R-C) interactions for depressive symptoms whilst controlling for residual trait correlations, including genetic correlation. The G-C and R-C interactions allow the polygenic variance component and the residual variance component for depressive symptoms respectively to vary across a continuous covariate trait. The MRNM extends existing polygenic-environment interaction approaches from categorical environmental traits to continuous ones, providing a route to avoid the pitfalls associated with the arbitrary categorising of continuous traits (Altman & Royston, 2006; Naggara, et al., 2011).

We included 25 continuous covariate traits, which covered childhood trauma, body composition, physical activity, smoking and PRSs for eight mental health disorders. For each covariate in turn, MRNMs with and without interactions were run and compared using LRTs. These models jointly test the presence of G-C and/or R-C interaction effects. The

contribution of each interaction is then extracted from the variance component estimates, summarised as the proportion of the outcome variability an interaction effect explains.

MRNMs for 11 of the 25 covariates found evidence for some interaction effect. Across these 11 covariates, the variability in depressive symptoms attributable to the R-C interaction effect tended to be substantially larger than that attributable to the G-C effect. Only one covariate, average sleep duration, had a G-C variance component estimate where the confidence interval excluded zero. The significant p-values observed in the LRTs between the null and interaction models are therefore likely to be driven by the R-C interactions. For nine covariate traits, the proportion of variability in depressive symptoms attributable to the R-C interaction had a confidence interval which did not include zero, but the proportion of variance accounted for was small, with the residual-'childhood trauma' interaction explaining the largest percentage of phenotypic variation at 3.0%, decreasing to 1.0% for the residual-'waist-to-hip ratio' interaction.

These R-C interactions can be interpreted as non-random noise, where the covariate traits explain additional variation in depSympt not captured explicitly in the model. Several possible extensions to the model may better explain the role of these covariate traits in depressive symptoms. Firstly, these variables may have a non-linear relationship with symptoms of depression not captured by fractional polynomials (FPs); refinement of the non-linear fixed effects model would resolve this e.g. fitting splines. Secondly, the covariates may interact with each other, or with additionally un-modelled environmental variables, to influence depressive symptoms. Finally, the covariates may interact with genetic variants not captured by genome-wide genotyping and imputation (such as rare variants, repeats, or structural variation), or with other omics-type data (such as the transcriptome).

Individually these R-C interactions are small, but cumulatively they could explain a large proportion of the variability in depressive symptoms. Random noise is not useful for prediction and further research to explain the residual heteroscedasticity is warranted. A potential route could be incorporating an environmental similarity matrix into the MRNM and looking for 'exposome' effects by utilising shared environmental information (Xuan, et al., 2020).

A further consideration is that the R-C interaction effect can capture deviations from normality in the conditional outcome trait; an effect not necessarily indicative of an interaction, rather driven by the mean-variance relationship of non-normal distributions (Young, Wauthier, & Donnelly, 2018). If this were true, then exploring the significant R-C interactions may not yield useful results. However, possible explanations for the significant R-C interactions, such as environment-environment interactions, have been reported in the depression literature (Hullam, et al., 2019; Morrissey & Kinderman, 2020), indicating that further exploration via multivariate approaches will improve the accuracy of depression models and reveal sets of relevant risk factors unlikely to be identified via univariate methods.

Our primary interest in these models was to assess evidence for G-C interactions for depressive symptoms. Only average sleep duration had an estimate of the proportion of phenotypic variability explained by a G-C interaction with a 95% confidence interval not

overlapping with zero. A small, but statistically significant, proportion of the variability in current depressive symptoms is therefore attributable to a genome-wide G-E interaction with average sleep duration measured at the UKB baseline assessment, 5-10 years earlier than the assessment of depSympt. Our results imply that the optimal sleep duration to minimise depressive symptoms can vary by genetic profile, but this modification of the polygenic variation for depSympt using historic sleep patterns is limited, with the estimate of variation attributable to this interaction being low at 1.2%. Other covariate traits that had non-zero estimates for R-C interaction variance component had much lower G-C interaction variance components, ranging from 0.59% (for childhood trauma), to estimates that were below zero (indicating a zero effect).

This is not the first study to identify a significant gene-sleep interaction for depression. A twin study by Watson, et al. (2014) found that the genetic contribution to depressive symptoms was significantly higher for both short (<7 hours per night) and long (≥ 9 hours per night) sleep durations compared to the average (7-8.9 hours per night)- a trend that we also observed (Supplementary Figure 46). Additionally, there is evidence for a complex bidirectional relationship between sleep and depression involving variables/ biomarkers such as circadian rhythms (Kronfeld-Schor & Einat, 2012; Khan, et al., 2018), stress (Leggett, Burgard, & Zivin, 2016; Palagini, et al., 2019), melatonin (Rahman, Marcu, Kayumov, & Shapiro, 2010), serotonin (van Dalsen & Markus, 2019), dopamine (Finan & Smith, 2013; Boland, Goldschmied, Wakschal, Nusslock, & Gehrman, 2020), and their respective genes. Future work investigating gene-sleep interactions for depression could utilise these previously highlighted genes within genomic partitioning analysis.

An initially surprising result is that this MRNM analysis does not support a non-zero G-C interaction effect for childhood trauma, despite G-C interactions of childhood trauma with depression previously having been identified (Peyrot, et al., 2014; Mullins, et al., 2016; Shen, et al., 2020). In addition to methodological and dataset differences, this study used a recently developed, composite measure of childhood trauma (Pitharouli, et al., 2021). The null G-C interaction result here, in contrast with previous significant results for other measures of childhood trauma, may suggest that G-C interactions for differing types of childhood trauma should be investigated as separate covariates and not as a continuous weighted aggregate.

Although estimated interaction effects for symptoms of depression are small, the MRNM has been able to identify strong lifestyle modulation effects on cardiovascular traits. Xuan et al. (2020) used the MRNM and 22 lifestyle traits to explore interactions for 23 cardiovascular traits, finding evidence of lifestyle modulation for 42% of the outcome-covariate trait pairings. Sizeable G-C and R-C interactions were observed suggesting the need for personalised lifestyle interventions to reduce the risk of cardiovascular disease. The largest G-C interaction explained $\sim 10\%$ of (fixed effects adjusted) phenotypic variability and was for the modulation of the polygenic effect on HDL cholesterol level by physical activity. The largest R-C interaction effect, explaining $\sim 20\%$ of phenotypic variability, was for the modulation of the nongenetic component for white blood cell count by smoking.

Our study has limitations. The environmental risk factors analysed here were selected based on a literature review, but no systematic review or meta-analysis was performed and other

variables with equally compelling rationale for inclusion may not have been considered. Similarly, PRSs for traits that have a genetic correlation with depression were included, together with biomarkers that have been widely tested for association with depression (CRP, vitamin D) but we cannot exclude interaction with other biomarkers. Additionally, the modifying effect of covariate traits on depressive symptoms were considered in separate models. A joint model for interactions may provide a better fit, however, the computational burden and the required sample size for robust parameter estimation will increase. All statistical models require assumptions about variables to be analysed. We analysed a continuous measure of depression, as required by the MRNM, and chose to use a composite measure previously constructed and validated as highly correlated with MDD diagnosis. Other options would have been to take raw scores for numbers of depression symptoms reported, or the ordinal measure of presence/absence of the two core depression symptoms (each scored 0, 1, 2), as analysed in other genome-wide association studies (Turley, et al., 2018; Levey, et al., 2021). The outcome trait *depSympt*, extracted from a hierarchical model, was chosen as it accounts for correlation between reported depression symptoms, and is continuous. The UKB depressive symptoms, from which the *depSympt* variable was constructed, represents a snapshot of participants mental health over the two weeks prior to completing the MHQ. This does not account for historical mental health, or capture trends of mental health, and the potential dynamic nature of G-E interactions, over time. Our study analysed only European ancestries in UKB and used only PRSs from GWASs of European ancestry. The MRNM provides a flexible and broad modelling framework, but model fitting is computationally intensive, particularly given the large sample sizes available in UKB. Meta-analysis across three subgroups was required to make this computational feasible. Ni et al. (2019) showed that using meta-analysis across subgroups reduced the power to identify non-zero interaction variance components compared to using the complete sample, and analysis with the full available UKB sample might yield more significant G-C/ R-C interactions. The MRNM used here estimates the polygenic-covariate interaction and, as an aggregate genome-wide measure, does not provide accessible information about the contribution from individual variants, genes, or pathways (Assary, Vincent, Machlitt-Northen, Keers, & Pluess, 2020), although the model could be extended for use within genomic partitioning.

In summary, the MRNM provides a flexible, if computationally intensive, framework to comprehensively model genetic and environmental contributions to complex traits. For depression, these models show significant R-C interactions, potentially highlighting unmodelled relationships between nongenetic contributions to depressive symptoms. Evidence for a G-C interaction was only found at one covariate (average sleep duration), suggesting that interactions between polygenic scores and the explored environmental risk variables do not play a major role in the etiology of depressive symptoms, and therefore, personalised lifestyle interventions based on SNP profiles are not required.

URLS

- LDSC HapMap 3 SNP-list: https://data.broadinstitute.org/alkesgroup/LDSCORE/w_hm3.snplist.bz2
- mtg2 software: <https://sites.google.com/site/honglee0707/mtg2>

Acknowledgements

This paper represents independent research funded by the UK Medical Research Council (MR/N015746/1), and the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The authors acknowledge use of the research computing facility at King's College London, Rosalind (<https://rosalind.kcl.ac.uk>), which is delivered in partnership with the NIHR Maudsley BRC, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR130505). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. Saskia P. Hagenaars was supported by the Medical Research Council (MR/S0151132). David M. Howard 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).

UKB: This research was conducted under UK Biobank application 18177.

Disclosures

Cathryn Lewis sits on the Myriad Neuroscience Scientific Advisory Board. The other authors declare no competing interests.

References

- Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080. doi:10.1136/bmj.332.7549.1080.
- Anney, R. J., Ripke, S., Anttila, V., Grove, J., Holmans, P., Huang, H., . . . Daly, M. J. (2017). Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism*, *8*, 21. doi:10.1186/s13229-017-0137-9.
- Arnau-Soler, A., Macdonald-Dunlop, E., Adams, M. J., Clarke, T.-K., MacIntyre, D. J., Milburn, K., . . . Thomson, P. A. (2019). Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. *Translational psychiatry*, *9*(1), 14. doi:10.1038/s41398-018-0360-y.
- Arnold, P. D., Askland, K. D., Barlassina, C., Bellodi, L., Bienvenu, O. J., Black, D., . . . OCD Collaborative Genetics Association Studies. (2018). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular Psychiatry*, *23*(5), 1181-1188. doi:10.1038/mp.2017.154.
- Assary, E., Vincent, J., Machlitt-Northen, S., Keers, R., & Pluess, M. (2020). The Role of Gene-Environment Interaction in Mental Health and Susceptibility to the Development of Psychiatric Disorders. Springer International Publishing.
- Baldwin, J. R., Reuben, A., Newbury, J. B., & Danese, A. (2019). Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment: A Systematic Review and Meta-analysis. *JAMA psychiatry*, *76*(6), 584-593. doi:10.1001/jamapsychiatry.2019.0097.
- Benner, A., & Ambler, G. (2015). *mfp: Multivariable Fractional Polynomials*. Retrieved from <https://CRAN.R-project.org/package=mfp>
- Boland, E. M., Goldschmied, J. R., Wakschal, E., Nusslock, R., & Gehrman, P. R. (2020). An Integrated Sleep and Reward Processing Model of Major Depressive Disorder. *Behavior Therapy*, *51*(4), 572-587. doi:10.1016/j.beth.2019.12.005.
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203-209. doi:10.1038/s41586-018-0579-z.
- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., . . . Poulton, R. (2003). Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene. *Science*, *301*(5631), 386. doi:10.1126/science.1083968.
- 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*(1). doi:10.1186/s13742-015-0047-8.
- Choi, K. W., Zheutlin, A. B., Karlson, R. A., Wang, M.-J., Dunn, E. C., Stein, M. B., . . . Smoller, J. W. (2020). Physical activity offsets genetic risk for incident depression assessed via electronic health records in a biobank cohort study. *Depression and Anxiety*, *37*(2), 106-114. doi:10.1002/da.22967.
- 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*(4), 298-304. doi:10.1093/bfgp/elv037.
- Coleman, J. R., Peyrot, W. J., Purves, K. L., Davis, K. A., Rayner, C., Choi, S. W., . . . Breen, G. (2020). Genome-wide gene-environment analyses of major depressive disorder and

- reported lifetime traumatic experiences in UK Biobank. *Molecular Psychiatry*, 25(7), 1430-1446. doi:10.1038/s41380-019-0546-6.
- Dahl, A., Nguyen, K., Cai, N., Gandal, M. J., Flint, J., & Zaitlen, N. (2020). A Robust Method Uncovers Significant Context-Specific Heritability in Diverse Complex Traits. *The American Journal of Human Genetics*, 106(1), 71-91. doi:10.1016/j.ajhg.2019.11.015.
- Davis, K. A., Coleman, J. R., Adams, M., Allen, N., Breen, G., Cullen, B., . . . Hotopf, M. (2020). Mental health in UK Biobank –development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open*, 6(2), e18. doi:10.1192/bjo.2019.100.
- Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., . . . Neale, B. M. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, 51(1), 63-75. doi:10.1038/s41588-018-0269-7.
- Dunn, E. C., Wiste, A., Radmanesh, F., Almlil, L. M., Gogarten, S. M., Sofer, T., . . . Smoller, J. W. (2016). Genome-wide association study (GWAS) and genome-wide by environment interaction study (GWEIS) of depressive symptoms in african american and hispanic/latina women. *Depression and anxiety*, 33(4), 265-280. doi:10.1002/da.22484.
- Evangelou, E., & Ioannidis, J. P. (2013). Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*, 14, 379-389. doi:10.1038/nrg3472.
- Finan, P. H., & Smith, M. T. (2013). The comorbidity of insomnia, chronic pain, and depression: Dopamine as a putative mechanism. *Sleep Medicine Reviews*, 17(3), 173-183. doi:10.1016/j.smrv.2012.03.003.
- Gianfredi, V., Blandi, L., Cacitti, S., Minelli, M., Signorelli, C., Amerio, A., & Odone, A. (2020). Depression and Objectively Measured Physical Activity: A Systematic Review and Meta-Analysis. *International Journal of Environmental Research and Public Health*, 17(10), 3738. doi:10.3390/ijerph17103738.
- Hullam, G., Antal, P., Petschner, P., Gonda, X., Bagdy, G., Deakin, B., & Juhasz, G. (2019). The UKB envirome of depression: from interactions to synergistic effects. *Scientific Reports*, 9(1), 9723. doi:10.1038/s41598-019-46001-5.
- Hunter, D. J. (2005). Gene–environment interactions in human diseases. *Nature Reviews Genetics*, 6(4), 287-298. doi:10.1038/nrg1578.
- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., . . . de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and applied genetics*, 127, 595-607. doi:10.1007/s00122-013-2243-1.
- Jermy, B. S., Hagenars, S. P., Glanville, K. P., Coleman, J. R., Howard, D. M., Breen, G., . . . Lewis, C. M. (2020). Using major depression polygenic risk scores to explore the depressive symptom continuum. *Psychological Medicine*, 1-10. doi:10.1017/S0033291720001828.
- Johnson, W., Mortensen, E. L., & Kyvik, K. O. (2020). Gene–Environment Interplay Between Physical Exercise and Fitness and Depression Symptomatology. *Behavior Genetics*, 50(5), 346-362. doi:10.1007/s10519-020-10009-9.
- Kendall, K. M., Van Assche, E., Andlauer, T. F., Choi, K. W., Luykx, J. J., Schulte, E. C., & Lu, Y. (2021). The genetic basis of major depression. *Psychological Medicine*, 1-14. doi:10.1017/S0033291721000441.

- Kerin, M., & Marchini, J. (2020). Inferring Gene-by-Environment Interactions with a Bayesian Whole-Genome Regression Model. *The American Journal of Human Genetics*, *107*(4), 698-713. doi:10.1016/j.ajhg.2020.08.009.
- Khan, S., Nabi, G., Yao, L., Siddique, R., Sajjad, W., Kumar, S., . . . Hou, H. (2018). Health risks associated with genetic alterations in internal clock system by external factors. *International journal of biological sciences*, *14*(7), 791-798. doi:10.7150/ijbs.23744.
- Kronfeld-Schor, N., & Einat, H. (2012). Circadian rhythms and depression: Human psychopathology and animal models. *Neuropharmacology*, *62*(1), 101-114. doi:10.1016/j.neuropharm.2011.08.020.
- Lee, S. H., & van der Werf, J. H. (2016). MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics*, *32*(9), 1420-1422. doi:10.1093/bioinformatics/btw012.
- Leggett, A., Burgard, S., & Zivin, K. (2016). The Impact of Sleep Disturbance on the Association Between Stressful Life Events and Depressive Symptoms. *The Journals of Gerontology: Series B*, *71*(1), 118-128. doi:10.1093/geronb/gbv072.
- Levey, D. F., Stein, M. B., Wendt, F. R., Pathak, G. A., Zhou, H., Aslan, M., . . . Gelernter, J. (2021). Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nature Neuroscience*, *24*(7), 954-963. doi:10.1038/s41593-021-00860-2.
- Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W., & Zitman, F. G. (2010). Overweight, Obesity, and Depression: A Systematic Review and Meta-analysis of Longitudinal Studies. *Archives of General Psychiatry*, *67*(3), 220-229. doi:10.1001/archgenpsychiatry.2010.2.
- Morrissey, K., & Kinderman, P. (2020). The impact of childhood socioeconomic status on depression and anxiety in adult life: Testing the accumulation, critical period and social mobility hypotheses. *SSM - Population Health*, *11*, 100576. doi:10.1016/j.ssmph.2020.100576.
- Mullins, N., Power, R. A., Fisher, H. L., Hanscombe, K. B., Euesden, J., Iniesta, R., . . . Lewis, C. M. (2016). Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychological medicine*, *46*(4), 759-770. doi:10.1017/S0033291715002172.
- Naggara, O., Raymond, J., Guilbert, F., Roy, D., Weill, A., & Altman, D. G. (2011). Analysis by Categorizing or Dichotomizing Continuous Variables Is Inadvisable: An Example from the Natural History of Unruptured Aneurysms. *American Journal of Neuroradiology*, *32*(3), 437. doi:10.3174/ajnr.A2425.
- Ni, G., van der Werf, J., Zhou, X., Hyppönen, E., Wray, N. R., & Lee, S. H. (2019). Genotype-covariate correlation and interaction disentangled by a whole-genome multivariate reaction norm model. *Nature Communications*, *10*(1), 2239. doi:10.1038/s41467-019-10128-w.
- Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., . . . Hettema, J. M. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular Psychiatry*, *21*(10), 1391-1399. doi:10.1038/mp.2015.197.
- Palagini, L., Domschke, K., Benedetti, F., Foster, R. G., Wulff, K., & Riemann, D. (2019). Developmental pathways towards mood disorders in adult life: Is there a role for sleep disturbances? *Journal of Affective Disorders*, *243*, 121-132. doi:10.1016/j.jad.2018.09.011.

- Peterson, R. E., Cai, N., Dahl, A. W., Bigdeli, T. B., Edwards, A. C., Webb, B. T., . . . Kendler, K. S. (2018). Molecular Genetic Analysis Subdivided by Adversity Exposure Suggests Etiologic Heterogeneity in Major Depression. *American Journal of Psychiatry*, *175*(6), 545-554. doi:10.1176/appi.ajp.2017.17060621.
- Peyrot, W. J., Milaneschi, Y., Abdellaoui, A., Sullivan, P. F., Hottenga, J. J., Boomsma, D. I., & Penninx, B. W. (2014). Effect of polygenic risk scores on depression in childhood trauma. *The British journal of psychiatry : the journal of mental science*, *205*(2), 113-119. doi:10.1192/bjp.bp.113.143081.
- Peyrot, W. J., Van der Auwera, S., Milaneschi, Y., Dolan, C. V., Madden, P. A., Sullivan, P. F., . . . Penninx, B. W. (2018). Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biological psychiatry*, *84*(2), 138-147. doi:10.1016/j.biopsych.2017.09.009.
- Pitharouli, M. C., Hagenaars, S. P., Glanville, K. P., Coleman, J. R., Hotopf, M., Lewis, C. M., & Pariante, C. M. (2021). Depressed patients have elevated C-reactive protein independently of genetic, health and psychosocial factors, in the UK Biobank. *American Journal of Psychiatry*, *178*(6), 522-529. doi:10.1176/appi.ajp.2020.20060947.
- Purcell, S. (2002). Variance Components Models for Gene–Environment Interaction in Twin Analysis. *Twin Research*, *5*(6), 554-571. doi:10.1375/136905202762342026.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing) Retrieved from <https://www.R-project.org/>
- Rahman, S. A., Marcu, S., Kayumov, L., & Shapiro, C. M. (2010). Altered sleep architecture and higher incidence of subsyndromal depression in low endogenous melatonin secretors. *European Archives of Psychiatry and Clinical Neuroscience*, *260*(4), 327-335. doi:10.1007/s00406-009-0080-7.
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., . . . Alexander, M. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421-427. doi:10.1038/nature13595.
- Royston, P., & Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling (with discussion). *Applied Statistics*, *43*(3), 429-467. doi:10.2307/2986270.
- Schaeffer, L. R. (2004). Application of random regression models in animal breeding. *Livestock Production Science*, *86*(1), 35-45. doi:10.1016/S0301-6226(03)00151-9.
- Shen, X., Howard, D. M., Adams, M. J., Hill, W. D., Clarke, T.-K., Consortium, M. D., . . . McIntosh, A. M. (2020). A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. *Nature communications*, *11*(1), 2301. doi:10.1038/s41467-020-16022-0.
- Sklar, P., Ripke, S., Scott, L. J., Andreassen, O. A., Cichon, S., Craddock, N., . . . Purcell, S. M. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics*, *43*(10), 977-983. doi:10.1038/ng.943.
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., . . . Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine*, *12*(3), e1001779. doi:10.1371/journal.pmed.1001779.

- Turley, P., Walters, R. K., Maghzian, O., Okbay, A., Lee, J. J., Fontana, M. A., . . . Benjamin, D. J. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics*, *50*(2), 229-237. doi:10.1038/s41588-017-0009-4.
- van Dalen, J. H., & Markus, C. R. (2019). The involvement of sleep in the relationship between the serotonin transporter gene-linked polymorphic region (5-HTTLPR) and depression: A systematic review. *Journal of Affective Disorders*, *256*, 205-212. doi:10.1016/j.jad.2019.05.047.
- Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., & Murray, C. J. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, *388*(10), 1545-1602. doi:10.1016/S0140-6736(16)31678-6.
- Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., . . . Murray, C. J. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, *396*(10258), 1204-1222. doi:10.1016/S0140-6736(20)30925-9.
- Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R., Gaspar, H. A., . . . Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, *51*(8), 1207-1214. doi:10.1038/s41588-019-0439-2.
- Watson, N. F., Harden, K. P., Buchwald, D., Vitiello, M. V., Pack, A., S. E., & Goldberg, J. (2014). Sleep duration and depressive symptoms: a gene-environment interaction. *Sleep*, *37*(2), 351-358. doi:10.5665/sleep.3412.
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., & Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*, *50*(5), 668-681. doi:10.1038/s41588-018-0090-3.
- Xuan, Z., Julius, v. d., Kristin, C.-C., Guiyan, N., John, M., Elina, H., & Hong, L. S. (2020). Whole-Genome Approach Discovers Novel Genetic and Nongenetic Variance Components Modulated by Lifestyle for Cardiovascular Health. *Journal of the American Heart Association*, *9*(8), e015661. doi:10.1161/JAHA.119.015661.
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American journal of human genetics*, *88*(1), 76-82. doi:10.1016/j.ajhg.2010.11.011.
- Young, A. I., Wauthier, F. L., & Donnelly, P. (2018). Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. *Nature Genetics*, *50*(11), 1608-1614. doi:10.1038/s41588-018-0225-6.

Tables

Covariate	p-value	Proportion of variability in depSympt [†] attributable to:			
		G-C interaction (%)		R-C interaction (%)	
		Estimate	95% CI	Estimate	95% CI
Neuroticism	5.06E-139	-0.15	[-0.76, 0.46]	2.58	[1.86, 3.30]
Childhood trauma	2.59E-058	0.59	[-0.14, 1.32]	2.98	[2.18, 3.77]
Sleep	1.97E-041	1.22	[0.54, 1.89]	2.52	[1.78, 3.27]
BMI	6.36E-018	-0.23	[-0.86, 0.41]	1.39	[0.68, 2.09]
Waist circumference	6.15E-016	-0.15	[-0.78, 0.48]	1.48	[0.78, 2.19]
Smoking	2.49E-010	0.47	[-0.52, 1.46]	1.57	[0.51, 2.63]
Waist-to-hip ratio	4.49E-009	-0.33	[-0.95, 0.29]	1.03	[0.34, 1.73]
MET total	1.92E-007	0.23	[-0.42, 0.87]	0.53	[-0.17, 1.24]
MET walk	1.13E-005	0.10	[-0.55, 0.74]	1.18	[0.45, 1.92]
MET mod	5.73E-004	-0.26	[-0.87, 0.35]	-0.08	[-0.78, 0.61]
TDI	1.96E-003	-0.19	[-0.81, 0.42]	1.67	[0.97, 2.38]

Table 1. Percentage of variation in depSympt (†adjusted for age, sex, batch and PCs) attributable to the genotype-covariate (G-C) interaction and the residual-covariate (R-C) interaction for covariate traits with significant interaction effects, showing p-value for comparison of the full model to null model, with significance set at $\alpha = 0.05/25 = 0.002$.

Figures

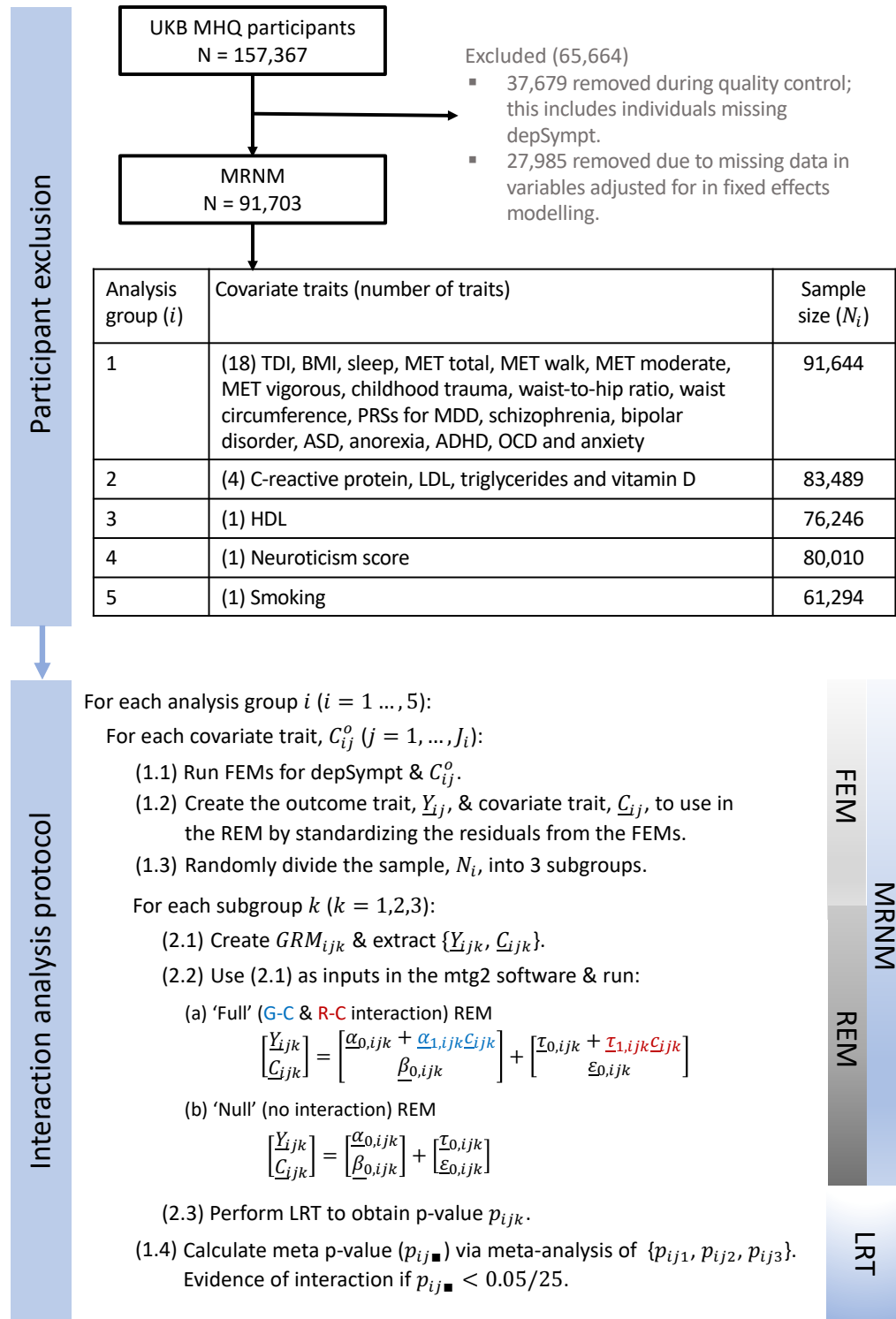


Figure 1. Methods flowchart for UK Biobank: participant exclusion, analysis groupings and scheme for interaction analysis. MRNM = multivariate reaction norm model. FEM = fixed effects model. REM = random effects model. G-C = genotype-covariate. R-C = residual-covariate. LRT = likelihood ratio test. N_{ik} = sample size for analysis group *i*, subgroup *k*. C_{ij} (C_{ijk}) = vector of length N_i (N_{ijk}) containing the *j*th standardised residual covariate trait from analysis group *i* (and subgroup *k*). Y_{ij} (Y_{ijk}) = vector of length N_i (N_{ijk}) containing the standardised residual outcome trait for interaction analysis with C_{ij} (C_{ijk}). $\alpha_{0,ijk}$ ($\beta_{0,ijk}$) = random effects vector of length N_{ik} representing the contribution to the outcome (covariate) trait for each individual from the homogeneous polygenic component. $\tau_{0,ijk}$ ($\epsilon_{0,ijk}$) = random effects vector of length N_{ik} representing the contribution to the outcome (covariate) trait for each individual from the homogeneous residual component). $\alpha_{1,ijk}$ ($\tau_{1,ijk}$) = random effects vector of length N_{ik} representing the contribution to the outcome trait for each individual due to G-C (R-C) interaction.

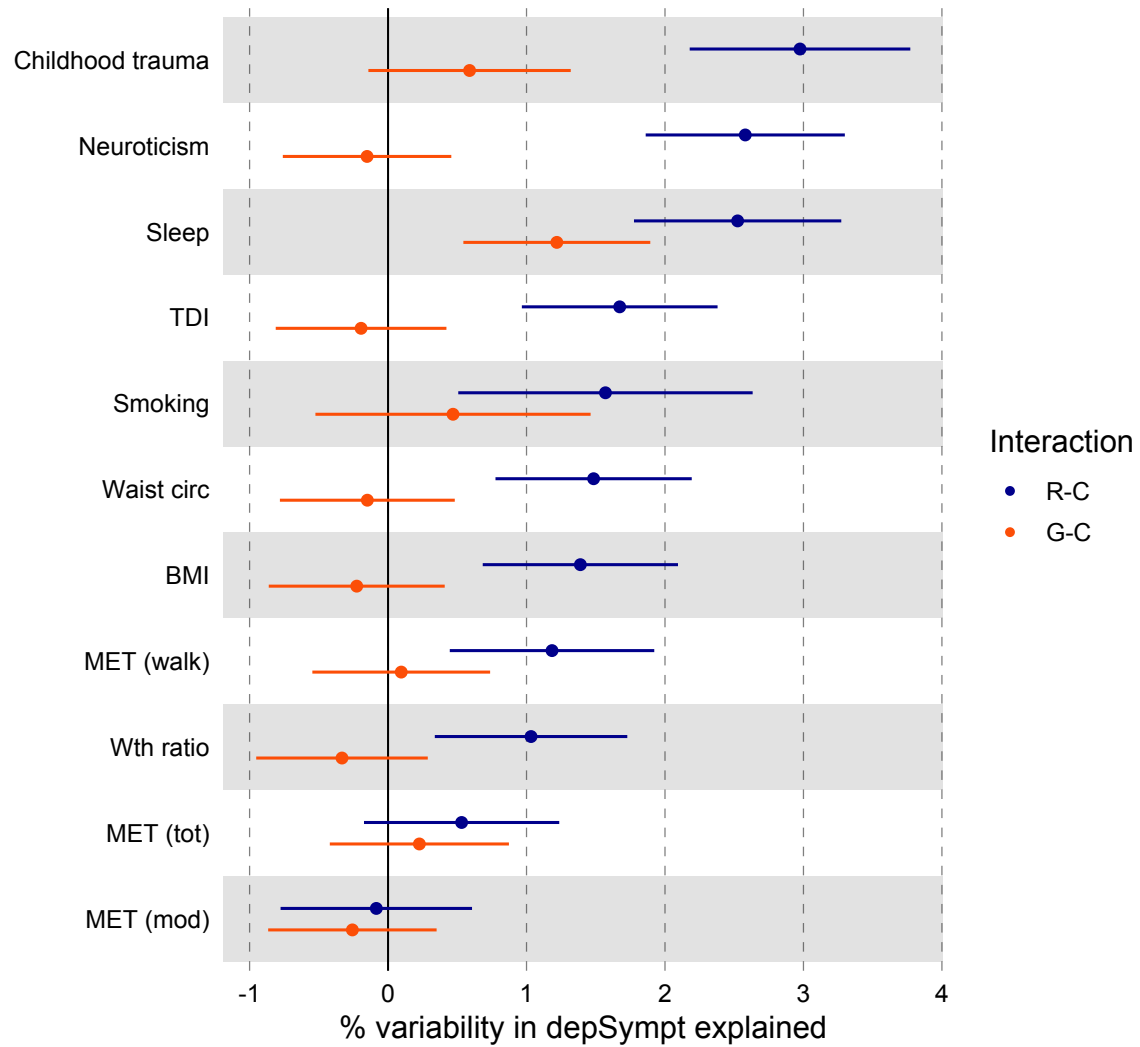


Figure 2. The percentage of variation in *depSympt* (adjusted for age, sex, batch and PCs) attributable to the R-C (residual-covariate) interaction (blue) and the G-C (genotype-covariate) interaction (red) with 95% confidence intervals.

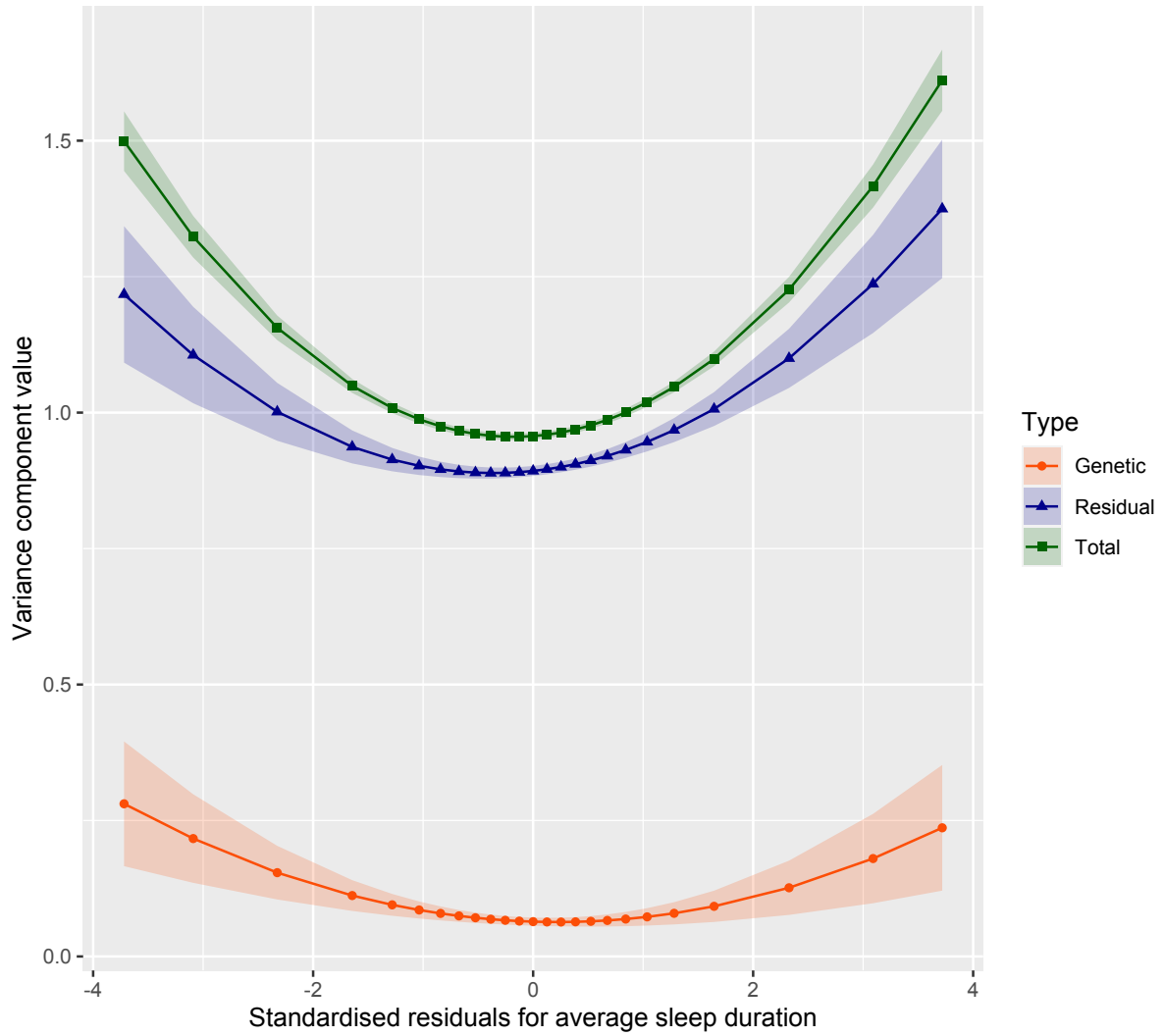


Figure 3. Variance components for standardised residual *depSympt* by standardised residual average sleep duration, with 95% confidence intervals (presented as coloured bands). See SM section 1.5.3 for variance component and standard error equations.