Cerebral Cortex, December 2020;30: 6121–6134

doi: 10.1093/cercor/bhaa158
Advance Access Publication Date: 17 July 2020
Original Article

ORIGINAL ARTICLE

Partial Support for an Interaction Between a Polygenic Risk Score for Major Depressive Disorder and Prenatal Maternal Depressive Symptoms on Infant Right Amygdalar Volumes

H. Acosta[1,2], K. Kantojärvi[3,4], N. Hashempour[1], J. Pelto[1], N. M. Scheinin[1,5], S. J. Lehtola[1], J. D. Lewis[6], V. S. Fonov[6], D. L. Collins[6], A. Evans[6], R. Parkkola[7], T. Lähdesmäki[8], J. Saunavaara[9], L. Karlsson[1,10], H. Merisaari[1,11,12], T. Paunio[3,4], H. Karlsson[1,5] and J. J. Tuulari[1,5,13,14]

[1]FinnBrain Birth Cohort Study, Turku Brain and Mind Center, Institute of Clinical Medicine, University of Turku, 20500 Turku, Finland, [2]Department of Psychiatry and Psychotherapy, Philipps University of Marburg, 35037 Marburg, Germany, [3]Finnish Institute for Health and Welfare, Genomics and Biobank Unit, FI-00271 Helsinki, Finland, [4]Department of Psychiatry and SleepWell Research Program, Faculty of Medicine, Helsinki University Central Hospital, University of Helsinki, 00100 Helsinki, Finland, [5]Department of Psychiatry, Turku University Hospital, University of Turku, 20500 Turku, Finland, [6]Montreal Neurological Institute, McGill University, Montreal H3A 0G4, Canada, [7]Department of Radiology, Turku University Hospital, University of Turku, 20500 Turku, Finland, [8]Department of Pediatric Neurology, Turku University Hospital, University of Turku, 20500 Turku, Finland, [9]Department of Medical Physics, Turku University Hospital, 20521 Turku, Finland, [10]Department of Child Psychiatry, Turku University Hospital, University of Turku, 20500 Turku, Finland, [11]Department of Future Technologies, University of Turku, 20500 Turku, Finland, [12]Center of Computational Imaging and Personalized Diagnostics, Case Western Reserve University, Cleveland, OH 44106, USA, [13]Turku Collegium for Science and Medicine, University of Turku, 20500 Turku, Finland and [14]Department of Psychiatry, University of Oxford, Oxford, OX1 2JD, UK

Address correspondence to Henriette Acosta, FinnBrain Birth Cohort Study, Turku Brain and Mind Center, Institute of Clinical Medicine, University of Turku, FI-20014 Turun yliopisto, 20500 Turku, Finland. Email: hschas@utu.fi, schneid5@staff.uni-marburg.de

Abstract

Psychiatric disease susceptibility partly originates prenatally and is shaped by an interplay of genetic and environmental risk factors. A recent study has provided preliminary evidence that an offspring polygenic risk score for major depressive disorder (PRS-MDD), based on European ancestry, interacts with prenatal maternal depressive symptoms (GxE) on neonatal right amygdalar (US and Asian cohort) and hippocampal volumes (Asian cohort). However, to date, this GxE interplay has only been addressed by one study and is yet unknown for a European ancestry sample. We investigated in 105 Finnish mother–infant dyads (44 female, 11–54 days old) how offspring PRS-MDD interacts with prenatal maternal depressive

symptoms (Edinburgh Postnatal Depression Scale, gestational weeks 14, 24, 34) on infant amygdalar and hippocampal volumes. We found a GxE effect on right amygdalar volumes, significant in the main analysis, but nonsignificant after multiple comparison correction and some of the control analyses, whose direction paralleled the US cohort findings. Additional exploratory analyses suggested a sex-specific GxE effect on right hippocampal volumes. Our study is the first to provide support, though statistically weak, for an interplay of offspring PRS-MDD and prenatal maternal depressive symptoms on infant limbic brain volumes in a cohort matched to the PRS-MDD discovery sample.

Key words: brain development, GxE interaction, hippocampus, limbic brain, MRI

Human brain structure is considered an intermediate phenotype for psychiatric diseases (Meyer-Lindenberg and Weinberger 2006). While numerous studies in adults have investigated how brain structural alterations are associated with psychiatric diseases and with related genetic risk factors, there is a dearth of comparable studies in infants. Therefore, little is known about how brain developmental trajectories from conception through infancy are affected by genetic risk factors for psychiatric diseases and by their interaction with the environment (Gao et al. 2019). Nevertheless, the origins of psychiatric disease susceptibility often trace back to early life stages, sometimes as early as the prenatal time period (Kim et al. 2015; O'Donnell and Meaney 2017). For instance, prenatal maternal distress has been shown to interact with genotype to increase the offspring's risk for psychiatric disorders later on (Rice et al. 2010). Thus, a better understanding of the etiology of psychiatric disorders such as major depressive disorder (MDD) should include an appreciation of the interplay of genetic variations with environmental factors on fetal brain development.

MDD is one of the leading causes for disability (Wittchen et al. 2011; Kyu et al. 2018) showing a moderately high heritability in the range of 31–42% (Sullivan et al. 2000). The heritability of MDD is likely accounted for by a high number of genetic variations with low effect sizes and manifold gene-by-gene and gene-by-environment interactions (in line with polygenic theory, first postulated for schizophrenia) (Gottesman and Shields 1967; Meyer-Lindenberg and Weinberger 2006; Demirkan et al. 2011). Recent genome-wide association studies (GWAS) identified genetic risk variants in individuals either diagnosed with MDD and/or with depressive symptoms (Ripke et al. 2013; Wray et al. 2018), and even though the GWAS mostly suffered from a lack of genome-wide significant results they are regarded as suitable for detecting MDD-related genes (Lubke et al. 2012). Taking into account the likely polygenic inheritance of MDD, the GWAS results have been used to establish polygenic risk scores for MDD (PRS-MDD): These risk scores are computed for each subject as the sum of risk alleles of single nucleotid polymorphisms (SNPs), which are selected from GWAS discovery samples up to a certain threshold of significance and weighted by their effect size. PRS-MDD have been shown to predict MDD risk (Peyrot et al. 2014; Taylor et al. 2019) and to interact with environmental risk factors for MDD, that is, childhood traumatic experiences (Peyrot et al. 2014).

As mentioned before, prenatal maternal distress constitutes an early environmental risk factor for child development (e.g., Rice et al. 2010). Prenatal maternal distress can be assessed as maternal depressive symptoms, and there is mounting evidence that prenatal maternal depressive symptoms are not only related to an increased risk for psychiatric disorders over the lifespan (O'Donnell and Meaney 2017) but also correlated with offspring's brain structural alterations, for example, in the right amygdala and frontal cortex (Rifkin-Graboi et al. 2013; Sandman et al. 2015; Lebel et al. 2016; Wen et al. 2017). However, to the best of our knowledge, to date only one study has probed the impact of the interaction of an offspring polygenic risk score for MDD with prenatal maternal depressive symptoms on infant brain structure (Qiu et al. 2017). Qiu et al. (2017) investigated mother–infant dyads of an Asian and a US cohort using a PRS-MDD from a comprehensive GWAS (Ripke et al. 2013). Qiu et al. (2017) reported a genotype-by-environment interaction (GxE) effect on infant right amygdalar volume in both cohorts (partly controlling for maternal genotype). In the US cohort, a significant positive association between prenatal maternal depressive symptoms of the third trimester and the offspring PRS-MDD (threshold $P = 0.1$) on right amygdalar volume was observed in neonates with low genetic risk (PRS-MDD < 0), and a marginal negative association was found in neonates with high genetic risk (PRS-MDD > 0). The interaction pattern was opposite in the Asian compared with the US sample. In the Asian cohort, but not the US cohort, an interaction effect of genotype with prenatal maternal depressive symptoms of the second trimester on right hippocampal volumes was additionally found. The difference in the results for the two cohorts was interpreted to stem from population differences between Asian and Caucasian samples affecting allele frequencies and the operation of risk alleles (Qiu et al. 2017).

The study by Qiu et al. (2017), as stated by the authors, was potentially limited by using a Caucasian sample as a discovery sample for studying the PRS-MDD in an Asian population. In the study of Qiu et al. (2017), the Asian sample consisted of individuals with a Chinese, Malay or Indian ethnic background. Moreover, the US sample, though referred to as Caucasian (Qiu et al., 2017), comprised Non-Hispanics (60%) and Hispanics (40%) of unspecified race or ancestry (i.e., persons with Native American, Black American and/or Caucasian descent). For instance, the individual ancestry of Hispanics in California, where the study has been performed, is 44% Native American and 8% African (Price et al. 2007). Hence, it seems unlikely that the ancestry of the US sample of Qiu et al. (2017) was predominantly of European descent. By contrast, the discovery sample (Ripke et al. 2013) of the PRS-MDD used in the study by Qiu et al. (2017) included only subjects of recent European ancestry (Ripke et al., 2013). The ancestry of a study population is pivotal for genetic studies given that average allele frequencies and risk genetic variants, such as for MDD, can differ between races and their geographic subpopulations (Lewontin and Hartl 1991; Porcelli et al. 2012). Accordingly, there is consistent evidence that polygenic risk scores derived from a predominantly European sample predict individual risk with higher accuracy in Europeans than in non-Europeans (Martin et al. 2018). Moreover, due to the challenges of imaging neonates, the sample sizes of both cohorts in the above mentioned study were modest (Asian cohort: $N = 168$, US cohort: $N = 85$) and, as acknowledged by the authors (Qiu et al. 2017), the findings preliminary.

Hence, the impact of the interaction of the offspring PRS-MDD (Ripke et al., 2013) with maternal prenatal depressive

symptoms on infant brain structure is still understudied. We aimed to address this gap in an independent European birth cohort. In more detail, we examined whether infant amygdalar and hippocampal volumes are significantly predicted by an interaction between offspring PRS-MDD and prenatal maternal depressive symptoms in Finnish mother–infant dyads. We assessed prenatal maternal depressive symptoms in the early and late second trimester and in the third trimester (gestational weeks 14, 24, and 34) and computed a sum score for the analyses. In post hoc analyses, we also investigated the interaction of genotype with maternal depressive symptoms of the different time points separately. Given the sexual dimorphism often observed for the effects of early stress on the amygdala (Buss et al. 2012; Wen et al. 2017; Soe et al. 2018) and early child development in general (Braithwaite et al. 2017; Braithwaite et al. 2018), we additionally explored sex-specific genetic, environmental, and GxE interaction effects.

Methods

Subjects

Participants were mother–infant dyads recruited from the FinnBrain Birth Cohort Study (www.finnbrain.fi) (Karlsson et al. 2018). Neuroimaging data were collected from 189 infants at the age of 1 to 8 weeks after birth. Mother–infant dyads were randomly selected according to the inclusion and exclusion criteria: inclusion criteria for neuroimaging were gestational age at birth ≥35 weeks and birth weight > 1500 g. Exclusion criteria were previously diagnosed central nervous system (CNS) anomalies or abnormal findings in a previous magnetic resonance imaging (MRI) scan. Written informed consent was obtained from all parent(s). The study was conducted according to the Declaration of Helsinki and was reviewed and approved by the Ethics Committee of the Hospital District of Southwest Finland (ETMK:31/180/2011). Of the 189 participants, 64 were excluded due to motion artifacts in the MR images. Of the remaining 125 infants, 19 were excluded because of missing genetic data (no GWAS performed [$N = 15$] or drop-out in QC [$N = 4$]). One more subject was excluded due to missing maternal Edinburgh Postnatal Depression Scale (EPDS) questionnaire data as described below. In the final sample, we included 105 infants (female: 41.9%, age after birth [days]: M = 26.1, SD = 7.2, range = 11–54, age of mothers [at term] [years]: M = 29.7 [SD = 4.6, range = 19–41]). The minimum birth weight of the included infants was 2530 g (missing: $N = 1$). Mothers did not significantly differ from the FinnBrain Birth Cohort Study mother population ($N = 2500$, Karlsson et al. 2018) with regard to EPDS sum scores ($t = -1.0$, $P = 0.324$) or education ($\chi^2 = 2.7$, $P = 0.257$), but they were on average 1.3 years younger ($t = 2.8$, $P = 0.005$).

Measures and Procedures

Prenatal Maternal Depressive Symptoms

Prenatal maternal depressive symptoms were assessed with the Finnish/Swedish version of the EPDS (Cox, J. Holden, J. Sagovsky 1987) at three time points during pregnancy (gestational weeks (gwk) 14, 24 and 34). Missing values at each time point (at maximum 3 items per time point) were imputed with the mean value of the existing ones. If no EPDS questionnaire data were available for one or two of the time points, data were imputed by the MissForest method (gwk14: $N = 2$, gwk24: $N = 1$, gwk34: $N = 6$) (Stekhoven and Bühlmann 2012). One mother did not provide any EPDS questionnaire data and was therefore excluded from the analyses. In this study, the individual sum scores of each trimester (EPDS gwk14, EPDS gwk24, EPDS gwk34) were investigated. Additionally, the individual sum scores of all three time points were combined to form a total EPDS sum score (EPDS sum). We observed outliers (> 3SD) (EPDS gwk14: $N = 3$, EPDS gwk24: $N = 2$, EPDS sum: $N = 1$) and excluded them in control analyses (see 2.2.6).

Other Maternal and Infant Variables

The following maternal variables were assessed via mothers' self-report at gwk 14 and/or 34: maternal education, maternal age, prenatal medication, and prenatal alcohol, nicotine and illicit drug consumption. Obstetric data were retrieved from the Finnish Medical Birth Register of the National Institute for Health and Welfare (http://www.thl.fi) and included infant's birth weight, gestational weeks, and apgar scores (apgar = appearance, pulse, grimace, activity, and respiration). We dichotomized maternal medication use as assessed at gwk14 (selective serotonin reuptake inhibitors [SSRI] and other CNS affecting medication, thyroxine, and corticosteroids; each yes/no) and alcohol, nicotine and/or illicit drug exposure (yes/no). Education was trichotomized [low: high school or vocational education (<12 years), middle: (career) college (12–15 years), high: university (+15 years)]. In the final sample 12 infants had a record of mild asphyxia.

MRI Acquisition

A detailed description of the MRI acquisition protocol is provided in a previous publication by the same research team (Lehtola et al. 2019). Participants were scanned with a Siemens Magnetom Verio 3 T scanner (Siemens Medical Solutions, Erlangen, Germany) during natural sleep. The 40-min imaging protocol included an axial PD-T2-TSE (Dual-Echo Turbo Spin Echo) sequence (repetition time [TR]: 12 070 ms, effective echo times [TE]: 13 ms and 102 ms) and a sagittal 3D-T1 MPRAGE (Magnetization Prepared Rapid Acquisition Gradient Echo) sequence (TR: 1900 ms, TE: 3.26, inversion time: 900 ms) with whole brain coverage and isotropic voxels of 1.0 mm^3 for both sequences. All brain images were assessed for incidental findings by a pediatric neuroradiologist.

Assessment of Structure Volumes

The volumes of the left and right amygdalae and hippocampi were assessed for each subject via label-fusion-based methods. These methods depend on achieving good registrations between the subjects and the template. This is increasingly difficult to achieve the further the template is from the subjects in terms of similarity. Thus, we constructed a template based on the subjects in this study. We then manually labeled the subcortical structures on this template, and constructed a library of warped versions of the labeled template such that the library represented the morphological variation in the sample. We then labeled the individual brains via label-fusion-based methods, and calculated the volume of each structure. The details of this approach are described below.

Construction of an unbiased population-specific template. All available, good quality imaging data ($N = 125$) were used for template construction. Each subject's T2 was linearly registered to their T1, and their T1 was linearly registered to the MNI 152 template. The average scaling from the native MRIs to the MNI 152 template was then computed, and the inverse used to scale the MNI 152 template to the average size of the population, which served as an initial target for construction of the population-specific

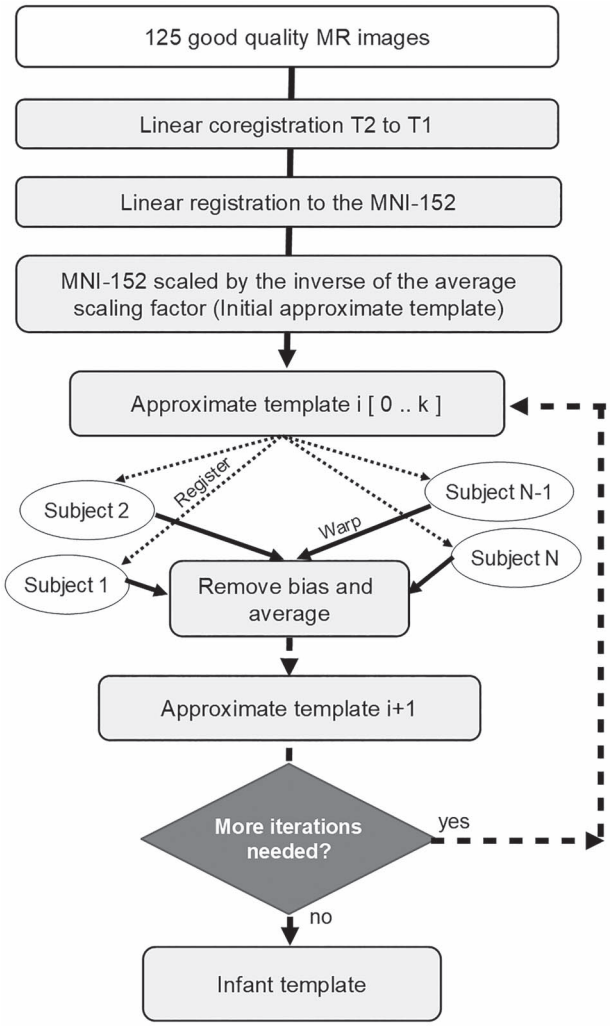

Figure 1. Iterative construction of the infant template as described in Fonov et al. (2011).

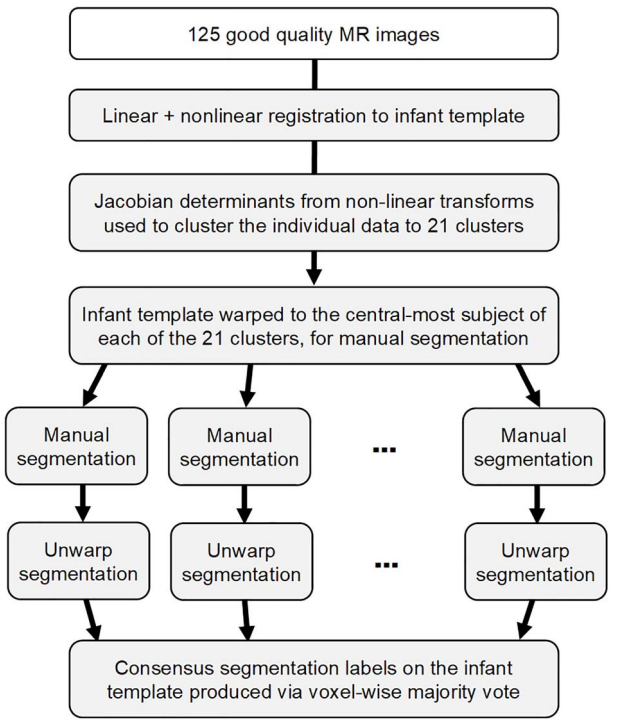

Figure 2. Labeling the infant template. The data were registered to the infant template, and then clustered based on the amount of distortion required to do that, into 21 clusters representing the morphological variability in the population. The template was then warped to the central-most subject of each cluster, providing 21 subtemplates for manual segmentation. After manual segmentation, the labels were then unwarped back to the base infant template, and merged via voxel-wise majority vote to create the consensus labels.

template, as described in Fonov et al. (2011). This iterative procedure builds a template that minimizes the mean squared intensity difference between the template and each subject's MRI and minimizes the magnitude of all deformations used to map the template to each subject's MRI. This method was applied to the T1 scans producing nonlinear transformations from the template space to each scan, then these transformations were used to map the T1 scans to the template space, where they were averaged to create the T1 template; these transformations were also combined with the T2 to T1 transformations to map the T2 scans into the common space where they were averaged to create the T2 template (Fig. 1).

Labeling the template. The structures of interest, that is, amygdalae, hippocampi, were manually labeled on the dual-contrast template. To ensure that these labels were accurate, we produced multiple variants of the template, and had each variant manually labeled (Fig. 2). The nonlinear transformations derived from the template construction procedure were used to cluster the subjects into 21 clusters from which 21 targets for manual segmentation of the subcortical structures were created. As the basis for clustering, the Jacobian was computed for the nonlinear transforms mapping each subject to the template. The values in the Jacobian were then extracted as a vector for each voxel within the template brain mask. These Jacobian vectors were then clustered using an equal combination of cosine similarity and Euclidean distance with Ward's clustering method (Ward 1963), with the number of clusters chosen to be 21. Then, within each cluster, the sum-squared distance from each subject to each other subject was computed, and the subject with the minimum sum-squared distance was taken as the central-most subject of the cluster. The dual-contrast template constructed in the previous step was then warped to these representative subjects ($N = 21$), and provided for manual segmentation (the persons performing the segmentation were unaware that these were, in fact, 21 different versions of the same template). The manual segmentations were then warped back to the standard template, and each voxel was assigned a label based on the majority vote across all 21 manual segmentations. This yielded the final labels for the amygdalae and hippocampi on the standard template (Fig. 2).

Manual segmentation. We used a previously published protocol for segmenting the bilateral amygdalae and hippocampi (Hashempour et al. 2019), noting that the 0.5 mm^3 resolution in the templates enabled more accurate segmentations. We started with one template segmented to serve as model by the primary rater NH and the senior rater JT and the segmentation was externally reviewed by JL. Then the final segmentations of the 21 subtemplates were performed and critically reviewed. The final labels are thus a consensus between the primary and the

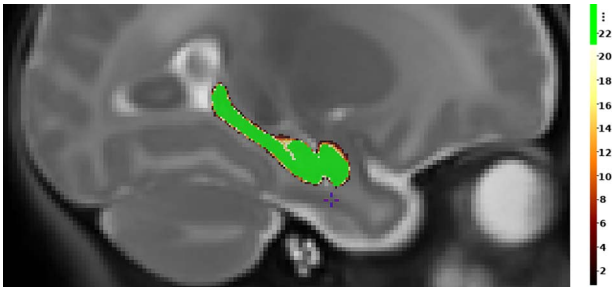

Figure 3. The sum of the binarized hippocampus and amygdala labels for each of the 21 templates and their x-flipped counterparts. The values show the degree to which all 42 manual segmentations agree (maximum = 42, minimum = 1, majority cut off = 22). The green color indicates labels with a majority agreement (≥22).

two senior raters. The base template was symmetric, and so the number of manual labelings was doubled by creating a copy that was flipped along the x-axis. The final labels were arrived at by majority agreement of both the flipped and unflipped manual segmentations of all 21 subtemplates (Fig. 3). This yielded labels for the hippocampus and amygdala that showed agreement across raters, hemispheres, and subtemplates as measured by the generalized conformity index (GCI). The hippocampus GCI was 0.763, and the amygdala GCI was 0.703. GCI scores of 0.7–1.0 are regarded as excellent (Kouwenhoven et al. 2009; Visser et al. 2019). The final majority agreement labels were then used for the segmentation of the individual images in the subsequent automated labeling step.

Labeling the subjects. Segmentation for each subject was done using a label-fusion-based labeling technique based on Coupé et al. (2011) and further developed by Weier et al. (2014) and by Lewis et al. (2019). The approach uses a population-specific template library for the hippocampi and amygdalae. The library was constructed by clustering (similarly to the method described above) the deformation fields from the nonlinear transforms produced during construction of the template and using the central-most subject of each cluster to construct the entries in the template library (Fig. 4). Thus, the template library represented the range of deformations found in the population. The clustering was done as described above, but using a dilated mask of the specific structures in order to capture the anatomical context of the nonlinear registration in that region of the brain, that is, in the region of hippocampus and amygdala, with the number of clusters now chosen as the square of the natural log of the number of subjects. The representative subject for each cluster was chosen as described above, that is, the central-most subject. This was done per hemisphere to accommodate hemispheric asymmetries, and the two sets of representative subjects were merged. To create the library entry for a cluster, the nonlinear transform for the central-most subject was used to warp the template together with the segmentation defined on it, and this pair was added to the template library. The template library was thus a set of warped copies of the template together with their correspondingly warped segmentations.

Once the template library had been created, each subject in the population was nonlinearly registered to the N closest templates in the library (here, $N = 7$), and the resulting transforms were used to warp their corresponding segmentations to the subject; the final labeling was then established via patch-based label fusion (Fig. 5). This was also done separately for each

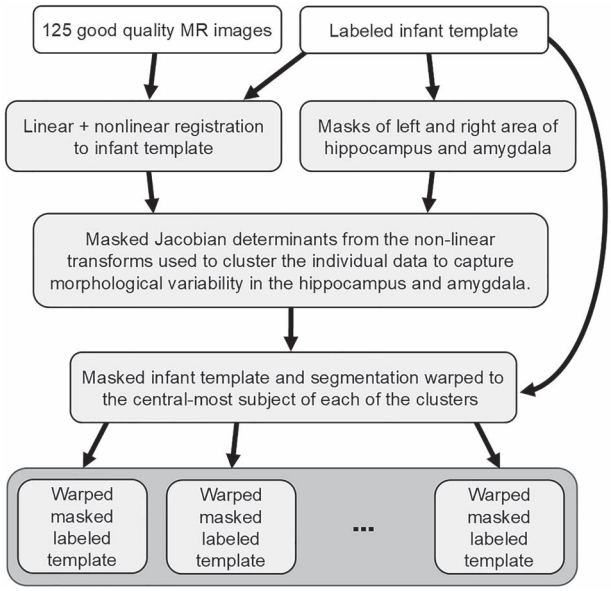

Figure 4. Creation of a template library. A mask of the hippocampus and amygdala was created and dilated to capture the anatomical context. The data were registered to the infant template, and the Jacobian of the nonlinear transform was masked, and used to cluster the subjects based on the distortion required to warp them to the template; the clusters thus represent the morphological variability in hippocampus and amygdala within the population. The template and its segmentation were then warped to the central-most subject of each cluster, and added to the library.

hemisphere. The volumes of each of the final labels were then computed and scaled to the native space based on the scaling factors in the subject's linear transforms, yielding the volumes of left and right hippocampus and amygdala (mm^3) for each subject. Similarly, a template space brain mask was scaled to native space and used to estimate total brain volume (TBV).

Genetic Analyses

An umbilical cord blood sample was drawn from each newborn at birth. DNA samples were extracted according to standard procedures at the National Institute for Health and Welfare and genotyped with Illumina Infinium PsychArray and Illumina Infinium Global Screening Array at the Estonian Genome Centre. QC was performed with PLINK 1.9 (www.cog-genomics.org/plink/1.9/) (Chang et al. 2015). Markers were removed for missingness (>5%) and Hardy-Weinberg equilibrium (P-value $< 1 \times 10^{-6}$). Individuals were checked for missing genotypes (>5%), relatedness (identical by descent calculation, PI_HAT>0.2) and population stratification (multidimensional scaling). Genotyped data were prephased with Eagle v2.4 (Loh et al. 2016) and imputed with Beagle v4.1 (Browning and Browning 2016) using the population-specific SISu v2 whole-genome sequencing data as imputation reference panel.

The individual offspring PRS-MDD were generated according to Qiu et al. (2017) and were based on a meta-analysis from the Psychiatric Genomics Consortium (discovery sample: European ancestry, 9240 MDD cases and 9519 controls) (Ripke et al. 2013). The PRS-MDD was computed as the sum of the number of risk alleles (0, 1, and 2) across SNPs weighted by their odds ratio as reported by Ripke et al. (2013) (retrieved from "https://www.med.unc.edu/pgc/results-and-downloads/downloads”). We selected the SNPs with the same threshold for

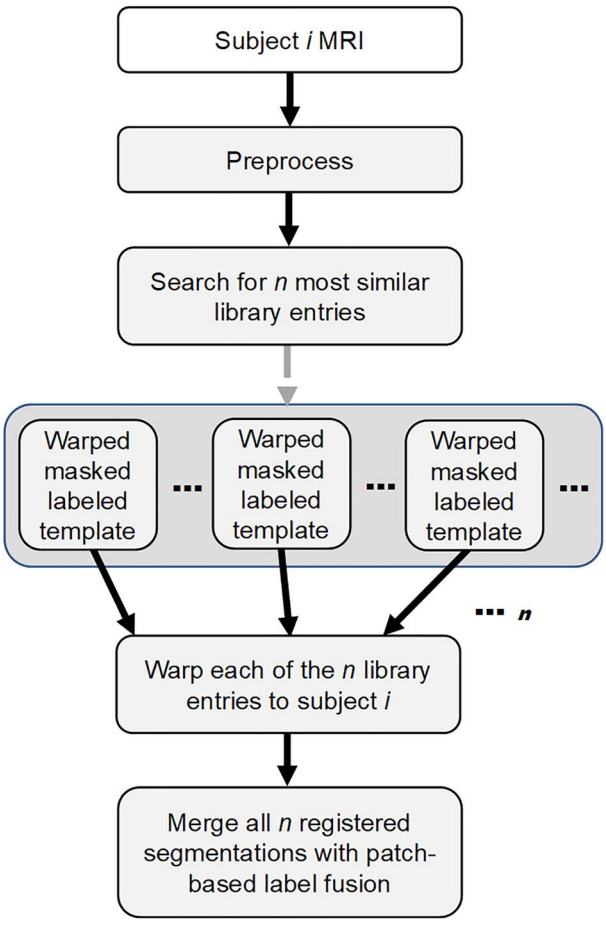

Figure 5. Automated labeling of a subject. The subject's MRI was preprocessed to produce a denoised, nonuniformity corrected version linearly aligned to the infant template. The N most similar library entries to the given subject were identified, masking for the context of the hippocampus and amygdala. Each of the N library entries were registered and warped to the subject, overlaying all N segmentations on the subject. Finally those N segmentations were merged via patch-based label fusion to create the labels for the hippocampus and amygdala for the subject.

linkage disequilibrium clumping as reported in Qiu et al. (2017) ($r^2 < 0.25$ within a 500 kb window) and which survived at *P*-value thresholds of 0.05, 0.1, and 0.2, respectively. Additionally, we investigated a PRS-MDD with SNPs that survived a *P*-value threshold of $P = 0.001$. Hence, the PRS-MDD consisted of 418, 12 878, 23 037, and 40 405 selected SNPs (due to imputation the numbers of selected SNPs differ from those reported in Qiu et al., 2017), and was standardized to a mean of zero for the infant MRI sample ($N = 161$).

Statistical Analyses

Statistical analyses were performed using R 3.4.4 (R Core Team 2016) (http://www.r-project.org/). Packages in use were "psych" (Revelle 2018), "nortest" (Gross and Ligges 2015), "ggplot2" (Wickham 2009) and "car" (Fox and Weisberg 2011) among others.

Standard multiple linear regression analyses were performed to probe the association of bilateral amygdalar and hippocampal volumes with individual genetic (PRS-MDD) and environmental (EPDS sum) risk scores. In post hoc analyses, we also probed the interaction between infant genotype and maternal EPDS scores of the second trimester (gwk 14 and 24) and third trimester (gwk 34) on bilateral amygdalar and hippocampal volumes.

In total, five models were applied to each outcome measure: First, the basic model, containing only the four control variables "infant's age after birth at MRI scan time," "gestational weeks," "total brain volume," and "infant's sex" (model B). These four control variables were included into every model. Second, the main effect of PRS-MDD (model G). Scores of the four different *P*-value thresholds were analyzed separately. Third, the main effect of EPDS (model E). Individual EPDS scores of the three time points and the EPDS sum score were investigated in independent analyses as described above. Fourth, an additive model of genetic and environmental effects (model G + E). Fifth, an interaction model of genetic and environmental effects (model GxE).

Two analysis of variance (ANOVA) model comparisons were applied to determine whether each model added explanatory value over the reduced model (see Halldorsdottir et al. 2019): in the first comparison, model B was compared with model G, model G was compared with model G + E and model G + E was compared with model GxE. The second comparison was done as the first one, but model G was replaced by model E. Models that significantly differed from the reduced model were further investigated in multiple regression analyses. We report the estimates, their standard errors (SEs) and *P*-values of significant predictors. Additionally, we report all results of GxE interaction models for bilateral hippocampal and amygdalar volumes with the predictors as reported in the US sample (EPDS of the third trimester, PRS-MDD $P = 0.1$) (Qiu et al., 2017).

In sensitivity analyses of significant results, we repeated the multiple regression analyses by subsequently adding and removing each of the following control variables to/from the model in order to test if the observed results were explained by these covariates: infant's birth weight, maternal education, maternal prenatal medication, and maternal prenatal alcohol, nicotine and/or illicit drug exposure. In the study of Qiu et al. (2017) mothers with psychotropic medication and infants with an apgar score < 9 had been excluded. Therefore, we conducted additional sensitivity analyses excluding all infants with an apgar score < 9 ($N = 14$, missings: $N = 1$, resulting sample: $N = 90$), as well as all mothers with prenatal medication (SSRI and other CNS affecting medication: $N = 9$, thyroxine: $N = 8$, glucocorticoids: $N = 6$, missings: $N = 4$, resulting sample: $N = 79$). In further sensitivity analyses we excluded EPDS outliers (> 3SD) (EPDS gwk14: $N = 3$, EPDS gwk24: $N = 2$, EPDS sum: $N = 1$) and infants with mild asphyxia at birth ($N = 12$, missings: $N = 2$, resulting sample: $N = 91$).

Finally, we explored sex-specific associations, that is, two-way interactions between sex and the genetic, or environmental scores, respectively, as well as three-way interactions between genetic scores, environmental scores, and sex. In ANOVA model comparisons, we compared the model of the sex-specific (S) interaction (model GxS, model ExS, model GxExS) against the model G, model E or model GxE, respectively. Only models that significantly differed from the reduced model were further investigated.

The significance threshold was set to $P < 0.05$. In order to control for the error rate related to multiple comparisons, we additionally report a false-discovery-rate (FDR) correction that was used for the four outcome measures (left/right amygdala, left/right hippocampus) (p.adjust function in R). No correction for multiple comparisons was applied for the different *P*-value thresholds of PRS-MDD, comparable to Qiu et al. (2017). We also did not apply multiple comparison corrections in our post hoc analyses of different EPDS time points and the exploratory

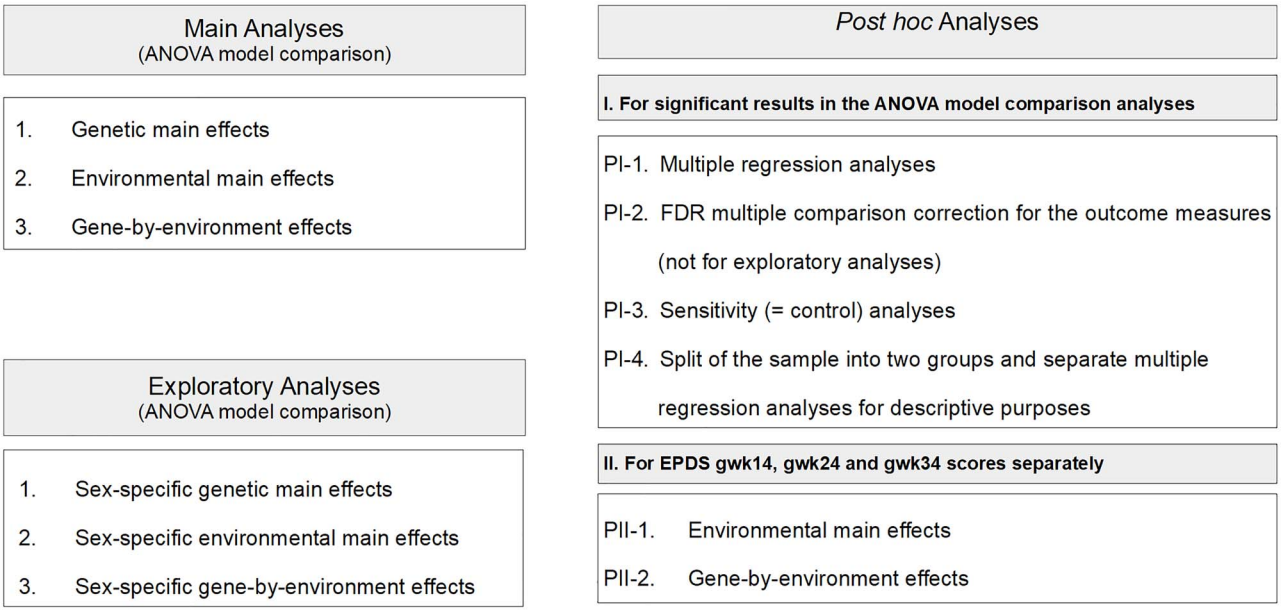

Figure 6. Schematic overview on the performed statistical analyses.

Table 1 The mean scores (*M*) and standard deviations (SD) or frequencies, respectively, are listed for maternal prenatal EPDS scores and control variables, for the whole sample (*N* = 105) and for girls (*N* = 44) and boys (*N* = 61) separately

Variable	Whole sample	Boys	Girls	
M ± SD (range)				
Child's age (days)	26.1 ± 7.2 (11–54)	26.9 ± 7.6 (11–43)	24.9 ± 6.5 (14–54)	0.153
Gestational weeks	39.8 ± 1.1 (36.3–42.1)	39.7 ± 1.0 (37.6–41.9)	40.0 ± 1.2 (36.3–42.1)	0.208
Birth weight (g) (*N* = 104)	3481.1 ± 418.8 (2530–4700)	3536.6 ± 438.2 (2720–4700)	3402.3 ± 380.8 (2530–4340)	0.108
Apgar, 5 min (*N* = 104)	9.0 ± 1.0 (4–10)	8.8 ± 1.1 (4–10)	9.2 ± 0.5 (8–10)	0.011*
EPDS (gwk 14)	5.51 ± 5.14 (0–25)	5.63 ± 4.99 (0–21)	5.36 ± 5.41 (0–25)	0.797
EPDS (gwk 24)	5.62 ± 5.28 (0–25)	5.99 ± 5.35 (0–23)	5.11 ± 5.22 (0–25)	0.404
EPDS (gwk 34)	5.51 ± 4.88 (0–20)	5.90 ± 5.06 (0–17)	4.98 ± 4.61 (0–20)	0.343
EPDS sum	16.65 ± 13.74 (0.00–63.95)	17.51 ± 13.88 (0.00–57.48)	15.45 ± 13.62 (1.00–63.95)	0.451
Frequencies				
Prenatal alcohol, nicotine and/or illicit drug consumption (no/yes) (*N* = 97)	64/33	37/21	27/12	0.579
Prenatal medication—SSRI and other CNS affecting drugs (no/yes) (*N* = 101)	92/9	52/7	40/2	0.217
Prenatal medication—thyroxine (no/yes) (*N* = 101)	93/8	53/6	40/2	0.321
Prenatal medication—corticosteroids (no/yes) (*N* = 101)	95/6	56/3	39/3	0.666
Maternal education (low/middle/high) (*N* = 103)	31/30/42	18/16/26	13/14/16	0.768

*$P < 0.05$. In the right column P-values for sex differences in the sample are listed.

analyses of sex-specific interactions. Figure 6 gives an overview on the performed statistical analyses.

Results

Demographic Overview

Demographic characteristics of the sample are presented in Table 1.

With regard to PRS-MDD, no significant sex-differences nor significant associations with maternal EPDS scores or control variables were observed. However, higher maternal EPDS scores were significantly associated with lower infant apgar scores (EPDS gwk 34: $r = -0.21$, $P = 0.036$, EPDS sum: $r = -0.20$, $P = 0.042$) and with maternal prenatal SSRI and/or other CNS affecting medication intake (EPDS gwk24: $t = -3.2$, $P = 0.002$, EPDS gwk34: $t = -3.1$, $P = 0.002$, EPDS sum: $t = -2.9$, $P = 0.005$). Significantly lower apgar scores were observed for male compared with

Table 2 Subcortical volumes and TBVs are listed for the whole sample and for boys and girls separately

Volumes (mm^3) (M ± SD)	Whole sample	Boys (N = 61)	Girls (N = 44)	P
Left amygdala	267.1 ± 37.8	272.3 ± 41.8	259.8 ± 30.5	0.093
Right amygdala	266.3 ± 39.3	277.1 ± 39.6	251.3 ± 34.0	<0.001
Left hippocampus	767.8 ± 116.1	793.0 ± 121.7	732.9 ± 98.9	0.008
Right hippocampus	768.4 ± 111.2	786.9 ± 109.5	742.6 ± 109.7	0.044
TBV	621791.5 ± 46752.9	633575.3 ± 45810.4	605454.8 ± 43454.8	0.002

In the right column P-values of sex differences in the sample are listed.

Table 3 The interaction effects of PRS-MDD (threshold: P = 0.1) and EPDS (all three time points and the sum score) on amygdalar and hippocampal volumes are shown as assessed in multiple regression analyses (estimates [β] with SE and P-values)

GxE	L Amygdala		R Amygdala		L Hippocampus		R Hippocampus	
PRS-MDD (P = 0.1) x	β ± SE	P (P[FDR])	β ± SE	P (P[FDR])	β ± SE	P (P[FDR])	β ± SE	P (P[FDR])
EPDS gwk14	−0.68 ± 0.71	0.338 [0.594]	−1.47 ± 0.73	0.046 [0.184]	1.15 ± 2.15	0.594 [0.549]	−1.38 ± 2.03	0.499 [0.594]
EPDS gwk24	−1.31 ± 0.76	0.090 [0.169]	−1.60 ± 0.79	0.047 [0.169]	0.20 ± 2.35	0.932 [0.932]	−3.37 ± 2.19	0.127 [0.169]
EPDS gwk34	−0.78 ± 0.80	0.332 [0.443]	−1.70 ± 0.82	0.041 [0.164]	−0.93 ± 2.43	0.702 [0.702]	−3.87 ± 2.26	0.090 [0.180]
EPDS sum	−0.37 ± 0.28	0.196 [0.261]	−0.66 ± 0.29	0.027 [0.108]	0.06 ± 0.87	0.947 [0.947]	−1.17 ± 0.81	0.153 [0.261]

female infants (Table 1). Female infants exhibited significantly smaller right amygdalar and bilateral hippocampal volumes than males, but also significantly smaller TBVs (Table 2).

Main analyses—Association of Infant Genotype and EPDS Sum Scores with Bilateral Amygdalar and Hippocampal Volumes

No Significant Main Genotype or Environmental Effects

The analyses revealed no significant main effect of infant genotype or of EPDS sum on bilateral amygdalar volumes, that is, model G (all $P > 0.14$) and model E (all $P > 0.39$) did not significantly add explanatory value to the reduced models in any of the ANOVA model comparisons. Similarly, no significant main effects were observed for bilateral hippocampal volumes (model E: all $P > 0.78$ and model G: all $P > 0.41$).

Significant Genotype-by-Environment Interaction Effects on Right Amygdalar Volumes

However, we observed that model GxE significantly differed from the additive model G + E for the right amygdalar volume, with PRS-MDD (threshold $P = 0.1$) and EPDS sum as predictors. The multiple regression analysis of this GxE interaction effect yielded that right amygdalar volumes were more positively associated with EPDS scores in infants with low PRS-MDD compared with infants with high PRS-MDD ($\beta = -0.66 \pm 0.29$, $P = 0.027$) (Table 3 and Fig. 7). However, this GxE effect did not survive the FDR correction for multiple comparisons ($P = 0.108$). No significant results were obtained for the PRS-MDD thresholds $P = 0.001$ ($F(97,1) = 0.4$, $P = 0.546$), $P = 0.05$ ($F(97,1) = 3.6$, $P = 0.061$) and $P = 0.2$ ($F(97,1) = 3.5$, $P = 0.066$) in the main analyses.

In the subsequent sensitivity analyses, the interaction effect of PRS-MDD and EPDS sum scores remained significant only after controlling for birth weight and EPDS outliers. In further post hoc analyses conducted as reported in Qiu et al. (2017), we split the sample into two groups according to the PRS-MDD, that is, we divided the sample into a group with infants with a low PRS-MDD (<0) and into one with a high PRS-MDD (>0). We tested the associations between EPDS sum scores and right amygdalar volumes in each group by means of regression analyses. Contrary to Qiu et al., (2017), who reported a significant positive association in the US cohort between maternal prenatal depressive symptoms and right amygdalar volumes in infants with low offspring PRS-MDD (<0), we found no significant associations in infants with a low PRS-MDD (<0) ($\beta = 0.65 \pm 0.50$, $P = 0.201$) or in infants with a high PRS-MDD (>0) ($\beta = -0.25 \pm 0.26$, $P = 0.336$).

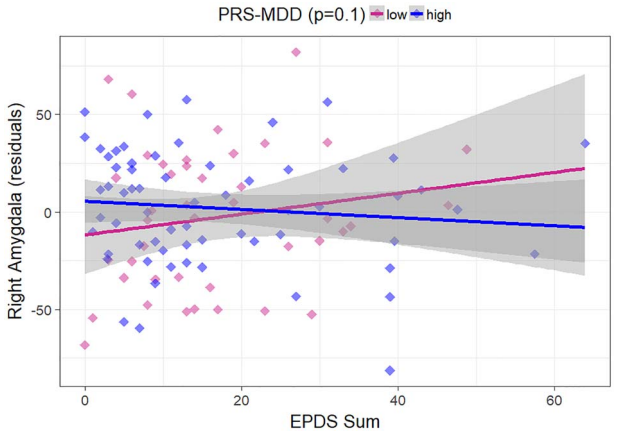

Figure 7. GxE effect on infant right amygdalar volumes. The interaction between offspring PRS-MDD (threshold: P = 0.1) and maternal EPDS sum scores on infant right amygdalar volumes (residuals, controlling for infant age, gestational weeks, TBV and sex) is depicted ($\beta = -0.66 \pm 0.29$, $P = 0.027$). Post hoc tests revealed that EPDS sum scores were weakly positively associated with right amygdalar volume in infants with a low PRS-MDD (<0) and weakly negatively in infants with a high PRS-MDD (>0) (the regression lines and their 95% confidence intervals are shown).

Significant GxE Interaction Effects on Right Amygdalar Volumes for the EPDS Scores of the Second and Third Trimester

In additional post hoc analyses (PII, see Fig. 6), we probed the association between infant genotype and EPDS scores of the second and third trimester (gwk14, gwk24, and gwk34) separately. We first investigated the main effects of the EPDS

scores of the second trimester (gwk 14 and gwk 24) and of the third trimester (gwk 34), and second its interaction with infant genotype for bilateral amygdalar and hippocampal volumes. We observed no significant main effects of EPDS gwk14 (all $P > 0.59$), EPDS gwk24 (all $P > 0.39$) and EPDS gwk34 scores (all $P > 0.31$). We observed significant GxE interaction effects on right amygdalar volumes with the PRS-MDD threshold of $P = 0.1$ and all EPDS measures (EPDS gwk14: $\beta = -1.47 \pm 0.73$., $P = 0.046$, EPDS gwk24: $\beta = -1.60 \pm 0.79$, $P = 0.047$, EPDS gwk 34: $\beta = -1.70 \pm 0.82$, $P = 0.041$, see also Table 3 and Supplementary Figure S1). No further significant GxE interaction effects were found for right amygdala (all $P > 0.06$), or left amygdalar or bilateral hippocampal volumes (all $P > 0.06$).

Exploratory analyses—Sexually Dimorphic Associations of Infant Genotype and EPDS Sum Scores with Amygdalar and Hippocampal Volumes

Significant Environment-by-Sex Interaction Effect on Right Amygdalar Volumes

Our exploratory analyses of sex differences yielded no significant genotype-by-sex interaction effects on amygdalar or hippocampal volumes (all $P > 0.16$), but we found that EPDS gwk24 scores were significantly more negatively associated with right amygdalar volumes in male compared with female infants ($\beta = -2.63 \pm 1.25$, $P = 0.038$, Supplementary Figure S2). This effect remained significant only after controlling for birth weight, maternal education and after excluding infants with mild asphyxia at birth. No further sex-specific associations of EPDS scores with the subcortical brain volumes were observed (all $P > 0.05$).

Significant Genotype-by-Environment-by-Sex Interaction Effects on Right Hippocampal Volumes

Exploring three-way interactions between PRS-MDD, EPDS scores and sex, we discovered a significant interaction effect on right hippocampal volumes: EPDS gwk24 scores showed a sex-specific interaction with PRS-MDD (all thresholds, strongest effects with $P = 0.1$) (PRS-MDD $P = 0.1$: $\beta = -14.52 \pm 4.26$, $P < 0.001$). This interaction effect stayed significant in all sensitivity analyses for PRS-MDD $P = 0.1$ (and $P = 0.001$).

Further post hoc analyses, splitting the sample into two groups, yielded no significant associations in infants with a low PRS-MDD (<0) (PRS-MDD $P = 0.1$) (male infants: $\beta = 5.07 \pm 3.38$, $P = 0.152$; female infants: $\beta = -4.55 \pm 5.77$, $P = 0.443$) or in infants with a high PRS-MDD (>0) (male infants: $\beta = -4.17 \pm 2.93$, $P = 0.164$, female infants: $\beta = 5.22 \pm 3.83$, $P = 0.188$) (Supplementary Figure S3). No further GxExS interaction effects were detected (all $P > 0.05$).

Discussion

A recent study has provided preliminary evidence that offspring PRS-MDD moderates the association between prenatal maternal depressive symptoms and the prenatal development of the right amygdala in an Asian and US cohort, and, in the Asian cohort, of the right hippocampus (Qiu et al. 2017). This study was potentially limited by studying a PRS-MDD, derived from a European ancestry discovery sample, in cohorts of (predominantly) non-European ancestry (Ripke et al. 2013; Qiu et al. 2017). To the best of our knowledge, our study is the first to probe the interaction between an offspring PRS-MDD and prenatal maternal depressive symptoms on infant limbic volumes in a European ancestry sample, in order to further elucidate the early developmental trajectory of psychiatric disease susceptibility.

In our study, no main effects of maternal depressive symptoms or of genotype on bilateral amygdalar or hippocampal volumes were found. However, we observed that offspring PRS-MDD (threshold $P = 0.1$) and prenatal maternal EPDS scores are jointly associated with infant right amygdalar volumes. This GxE effect was significant in the main analysis, but did not survive all control analyses, nor the correction for multiple comparisons. No GxE interaction effects on left amygdalar or bilateral hippocampal volumes were yielded. Post hoc analyses showed that right amygdalar volumes were more positively associated with EPDS scores of the second and third trimester (sum score and separate measures) in infants with low PRS-MDD compared with infants with high PRS-MDD. Hence, the direction of this GxE effect clearly resembled the one observed in the US cohort (Qiu et al., 2017). Thereby, our study results provide support, though statistically weak, for an interaction of genetic and environmental risk factors for MDD on neonatal amygdalar brain development which might affect offspring vulnerability for developing a psychiatric disorder later in life, such as MDD.

Exploring sex differences, we did not find significant sex-specific genotype effects on hippocampal or amygdalar volumes, but we observed evidence for sexually dimorphic associations with EPDS scores of the second trimester (gwk24): EPDS gwk24 scores were more negatively associated with right amygdalar volumes in male infants compared with female infants. Furthermore, we detected a sex-specific GxE effect on right hippocampal volumes: EPDS scores (gwk24) were more positively associated with right hippocampal volumes in male, genetic low risk infants, and more negatively in male genetic high risk infants compared with their female counterparts.

MDD is characterized by a biased attention, processing, and memory of emotional stimuli favoring negative over positive stimuli contrary to healthy controls (Disner et al. 2011). On a neural level, corticolimbic structures, such as the amygdala and the hippocampus, and cortico-striatal-pallidal-thalamic brain circuits have been implicated in the onset and course of depression (Drevets et al. 2008; Bora et al. 2012; Lindquist et al. 2012; Yilmazer-Hanke 2012). The amygdala, playing a key role in emotion processing, is presumably involved in the MDD-related bias for negative stimuli showing a stronger and longer lasting reactivity to negative emotional (e.g., fearful or sad) stimuli in adults (Siegle et al. 2002; Hamilton and Gottlib 2008; Peluso et al. 2009; Disner et al. 2011), adolescents (Roberson-Nay et al. 2006; Beesdo et al. 2009) and children (Gaffrey et al. 2011; Barch et al. 2012; Pagliaccio et al. 2012), in many, but not all (sub)studies (e.g., Thomas et al. 2001; Beesdo et al. 2009; Townsend et al. 2010). While the glucose metabolism in the amygdala has been shown to correlate positively with depressive symptom severity in adults (Drevets et al. 2008), meta-analyses yielded inconsistent associations of MDD with amygdalar volumes in adults (Hajek et al. 2009; Arnone et al. 2012) and in children/adolescents (Hajek et al. 2009). In infants, a larger right amygdalar volume predicted lower impulse control at 2 years of age (Graham et al. 2018). In children, larger bilateral amygdala volumes predicted decreased emotion regulation skills (Pagliaccio et al. 2014), and in girls a larger right amygdala volume was related to more affective problems (Buss et al. 2012) and a higher fearfulness (van der Plas et al. 2010). Larger amygdalar volumes have also been associated with an extended period of early adverse rearing conditions (i.e., institutional rearing) and were positively related to child's internalizing behavior and anxiety (Tottenham

et al. 2010). However, conflicting findings also exist, showing an association of larger amygdalar volumes with less emotional symptoms, less peer relationship problems (Acosta et al. 2019) and with less proactive aggression in children and adolescents (Naaijen et al. 2018). Interestingly, the observed GxE effect on right amygdalar volume suggests that larger right amygdalar volumes are only found when one risk factor, either environmental or genetic, is present, but in the case of both risk factors (i.e., higher infant genetic risk and higher maternal prenatal EPDS scores) a smaller right amygdalar volume is observed. Smaller amygdalar volumes have also been shown in hypercortisolic children and adults, albeit predominantly in the left hemisphere (Merke et al. 2005; Brown et al. 2008). Therefore, it is possible that the existence of both environmental and genetic risk factors—corresponding to a cumulation of adversity—leads to an hypercortisolic state and subsequently to a smaller amygdalar volume. However, as mentioned above, smaller amygdalar volumes have also been related to less fearfulness, better emotion regulation skills, more proactive aggression, and a higher impulse control which might be more adaptive to some postnatal environments. It has been put forward that early adversity during certain time windows can trigger adaptive responses in the genetically susceptible offspring enhancing the chances of survival in a specific environment. According to this hypothesis, a mismatch between the quality of early and later environment rather than adversity itself would put a genetically susceptible individual at risk (Nederhof and Schmidt 2012).

Our tentative finding on a sex-specific association of EPDS gwk24 scores with right amygdalar volumes is in line with other studies showing a more positive association of prenatal maternal distress on amygdalar volumes in girls (e.g., Buss et al. 2012; Wen et al. 2017; Acosta et al. 2019). However, conflicting findings exist in neonates: Qiu et al. (2017) and Rifkin-Graboi et al. (2013) did not report sexually dimorphic associations with amygdalar volumes. In a study overlapping with our sample, Lehtola et al. (n.d., unpublished data) found a sex-specific association of amygdalar volumes with prenatal stress (measured as a combination of maternal anxiety and depressive symptoms), with smaller amygdalar volumes in male compared with female infants. The observed sex differences might be due to sexually dimorphic placental functions in response to prenatal stress mediated by sex-specific placental epigenetic regulations, gene expressions, and receptor distributions (Clifton 2010; Bock et al. 2015). Sexually dimorphic neurodevelopmental trajectories might also underly the occurrence of sex differences (Bock et al. 2015; Entringer et al. 2015).

The MDD-related bias for negative stimuli involves hippocampal reactivity and volumes. The hippocampus plays a pivotal role in memory encoding and retrieval, and in stress physiology (Insausti and Amaral 2012). Higher hippocampal reactivity to negative versus neutral stimuli has been associated with depressive symptom severity in children (Pagliaccio et al. 2012), and with smaller child's hippocampal volumes (Suzuki et al. 2013). Reductions in hippocampal volumes have consistently been reported in persons with MDD irrespective of age (Arnone et al. 2012). While the extent of hippocampal volume shrinkage seems to correspond to the duration of illness (McKinnon et al. 2009), a smaller hippocampal volume also likely represents a vulnerability marker for MDD as the findings of a longitudinal study in adolescents suggest (Rao et al. 2010). Furthermore, animal studies demonstrated that dendritic atrophy in the hippocampus is related to chronic stress experiences (Vyas et al. 2002). Nevertheless, no main effects of maternal prenatal depressive symptoms on infant hippocampal volumes, but interaction effects with genetic factors on right hippocampal volumes have thus far been reported (Qiu et al. 2017; Wang et al. 2018). Accordingly, we did not observe main effects of maternal depressive symptoms or of infant genotype on hippocampal volumes in our study, but a sex-specific GxE interaction effect: we detected that right hippocampal volumes were more negatively associated with EPDS scores of the late second trimester in the male compared with the female genetic high risk group, suggesting a higher vulnerability to prenatal depressive symptoms (gwk24) in male, but not female genetic high risk infants. As mentioned above, sexually dimorphic neurodevelopmental trajectories and placental functions might account for the observed sex differences. However, given the modest sample size of our study, this three-way interaction effect has to be considered as tentative.

Limitations

The sample size of our study was marginally larger than the one of the US sample in the study of Qiu et al. (2017) ($N = 105$ compared with $N = 85$), but our findings were partly reduced to nonsignificance by the control analyses which might at least in parts be attributed to a lack of statistical power. The replication of true effects usually demands larger sample sizes than in the original study to reach adequate statistical power (Button et al. 2013). Hence, our study might still be underpowered. Future replication studies and meta-analyses are warranted to further corroborate the GxE interaction effect on infant right amygdalar volumes in a population of European ancestry (Lieberman and Cunningham 2009). However, a strength of our study was the genetic homogeneity given that all participants were native Finns.

Qiu et al. (2017) controlled for maternal genotype in the Asian, but not the US cohort. We had no information about maternal genotype at hand for our analyses. Results in the study of Qiu et al. (2017) stayed significant after controlling for maternal genotype, but we cannot rule out that maternal genotype interacts with prenatal maternal depressive symptoms on the maternal-placental-fetal stress physiology.

Furthermore, it would be worthwhile to investigate how the observed brain volume alterations are related to offspring emotion regulation skills and behavior later in development and how they interact with the postnatal environment to predict psychiatric disease susceptibility over the life course.

Conclusions

In sum, our data provide partial support that PRS-MDD and prenatal maternal EPDS scores of the second and third trimester jointly shape infant right amygdalar volumes, resembling the results of the US cohort of a recent study. Our study was the first to address this GxE interplay in a European ancestry sample that was matched to the discovery sample of the PRS-MDD. Additionally, EPDS scores of the second trimester were shown to interact with PRS-MDD on infant right hippocampal volumes in a sex-specific manner suggesting a higher vulnerability to environmental prenatal adversity in male infants with high PRS-MDD. Future studies should elucidate how the observed volume alterations of brain gray matter structures predict offspring behavior and psychiatric disease susceptibility in interaction with postnatal environments of various adversity.

Funding

This work was supported by the Jane and Aatos Erkko Foundation (to H.A. and H.K.); the Academy of Finland (grant numbers 264363, 253270, 134950 to H.K.; grant numbers 1350941 and 253346 to T.P.; grant number 308176 to L.K.); the Hospital District of Southwest Finland State Research Grants (grant number P3006 to J.T., grant number P3003 to NMS, grant number P3498 to H.K., grant number P3654 to L.K., K3562 to R.P.); the Signe and Ane Gyllenberg Foundation (to H.K., N.M.S., and L.K.); the Yrjö Jahnsson Foundation (grant number 6847 to L.K.); the Alfred Kordellin Foundation (to J.T.); the Turku University Foundation (to J.T.); the Emil Aaltonen Foundation (to J.T.); the Maire Taponen Foundation (to S.L.); the Juho Vainio Foundation (to S.L.); the Sigrid Jusélius Foundation (to J.T.); the Orion Research Foundation (to H.M. and J.T.); the NARSAD Brain and Behavior Research Foundation (grant number 1956 to L.K.); the Foundation for Pediatric Research (to R.P.); the Azrieli Neurodevelopmental Research Program (grant number ANRP-MIRI13–3388 to J.L. and A.E.); the Brain Canada Multi-Investigator Research Initiative (to A.E.); the Canadian Institutes of Health Research (to V.F. and D.C.); and the Natural Sciences and Engineering Research Council of Canada (to D.C.). The research also benefited from computational resources provided by Compute Canada (www.computecanada.ca) and Calcul Quebec (www.calculquebec.ca).

Conflict of Interest None declared.

References

Acosta H, Tuulari JJ, Scheinin NM, Hashempour N, Rajasilta O, Lavonius TI, Pelto J, Saunavaara V, Parkkola R, Lähdesmäki T *et al*. 2019. Maternal pregnancy-related anxiety is associated with sexually dimorphic alterations in amygdala volume in four-year-old children. *Front Behav Neurosci*. 13:175. doi: 10.3389/FNBEH.2019.00175.

Arnone D, McIntosh AM, Ebmeier KP, Munafò MR, Anderson IM. 2012. Magnetic resonance imaging studies in unipolar depression: systematic review and meta-regression analyses. *Eur Neuropsychopharmacol*. 22(1):1–16. doi: 10.1016/j.euroneuro.2011.05.003.

Barch DM, Gaffrey MS, Botteron KN, Belden AC, Luby JL. 2012. Functional brain activation to emotionally valenced faces in school-aged children with a history of preschool-onset major depression. *Biol Psychiatry*. 72(12):1035–1042. doi: 10.1016/j.biopsych.2012.06.009.

Beesdo K, Lau JYFF, Guyer AE, McClure-Tone EB, Monk CS, Nelson EE, Fromm SJ, Goldwin MA, Wittchen HU, Leibenluft E *et al*. 2009. Common and distinct amygdala-function perturbations in depressed vs anxious adolescents. *Arch Gen Psychiatry*. 66(3):275–285. doi: 10.1001/archgenpsychiatry.2008.545.

Bock J, Wainstock T, Braun K, Segal M. 2015. Stress in utero: prenatal programming of brain plasticity and cognition. *Biol Psychiatry*. 78(5):315–326. doi: 10.1016/j.biopsych.2015.02.036.

Bora E, Harrison BJ, Davey CG, Yücel M, Pantelis C. 2012. Meta-analysis of volumetric abnormalities in cortico-striatal-pallidal-thalamic circuits in major depressive disorder. *Psychol Med*. 42(4):671–681. doi: 10.1017/S0033291711001668.

Braithwaite EC, Hill J, Pickles A, Glover V, O'Donnell K, Sharp H. 2018. Associations between maternal prenatal cortisol and fetal growth are specific to infant sex: findings from the Wirral child health and development study. *J Dev Orig Health Dis*. 9(4):425–431. doi: 10.1017/S2040174418000181.

Braithwaite EC, Pickles A, Sharp H, Glover V, O'Donnell KJ, Tibu F, Hill J. 2017. Maternal prenatal cortisol predicts infant negative emotionality in a sex-dependent manner. *Physiol Behav*. 175:31–36. doi: 10.1016/j.physbeh.2017.03.017.

Brown ES, Woolston DJ, Frol AB. 2008. Amygdala volume in patients receiving chronic corticosteroid therapy. *Biol Psychiatry*. 63(7):705–709. doi: 10.1016/j.biopsych.2007.09.014.

Browning BL, Browning SR. 2016. Genotype imputation with millions of reference samples. *Am J Hum Genet*. 98(1):116–126. doi: 10.1016/j.ajhg.2015.11.020.

Buss C, Davis EP, Shahbaba B, Pruessner JC, Head K, Sandman CA. 2012. Maternal cortisol over the course of pregnancy and subsequent child amygdala and hippocampus volumes and affective problems. *Proceedings of the National Academy of Sciences of the United States of America*. 109(20):E1312–E1319. doi: 10.1073/pnas.1201295109.

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR. 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 14(5):365–376. doi: 10.1038/nrn3475.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 4(1). doi: 10.1186/s13742-015-0047-8.

Clifton VL. 2010. Review: sex and the human placenta: mediating differential strategies of fetal growth and survival. *Placenta*. 31(SUPPL.):S33–S39. doi: 10.1016/j.placenta.2009.11.010.

Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, Collins DL. 2011. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *Neuroimage*. 54(2):940–954. doi: 10.1016/j.neuroimage.2010.09.018.

Cox J, Holden J, Sagovsky R. 1987. Detection of postnatal depression: development of the 10-item Edinburgh postnatal depression scale. *Br J Psychiatry*. 150(6): 782–786.

Demirkan A, Penninx BWJH, Hek K, Wray NR, Amin N, Aulchenko YS, Van Dyck R, De Geus EJC, Hofman A, Uitterlinden AG *et al*. 2011. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol Psychiatry*. 16(7):773–783. doi: 10.1038/mp.2010.65.

Disner SG, Beevers CG, Haigh EAP, Beck AT. 2011. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci*. 12(8):467–477. doi: 10.1038/nrn3027.

Drevets WC, Price JL, Furey ML. 2008. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain Struct Funct*. 213(1–2):93–118. doi: 10.1007/s00429-008-0189-x.

Entringer S, Buss C, Wadhwa PD. 2015. Prenatal stress, development, health and disease risk: a psychobiological perspective −2015 Curt Richter award winner. *Psychoneuroendocrinology*. 62:366–375. doi: 10.1016/j.psyneuen.2015.08.019.

Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. 2011. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*. 54(1):313–327. doi: 10.1016/J.NEUROIMAGE.2010.07.033.

Fox J, Weisberg S. 2011. *An {R} Companion to Applied Regression*. 2nd ed. Thousand Oaks CA: Sage.

Gaffrey MS, Luby JL, Belden AC, Hirshberg JS, Volsch J, Barch DM. 2011. Association between depression severity and amygdala reactivity during sad face viewing in depressed preschoolers: an fMRI study. *J Affect Disord*. 129(1–3):364–370. doi: 10.1016/J.JAD.2010.08.031.

Gao W, Grewen K, Knickmeyer RC, Qiu A, Salzwedel A, Lin W, Gilmore JH. 2019. A review on neuroimaging studies of genetic and environmental influences on early brain development. *Neuroimage*. 185:802–812. doi: 10.1016/j.neuroimage.2018.04.032.

Gottesman II, Shields J. 1967. A polygenic theory of schizophrenia. *Proc Natl Acad Sci*. 58:199–205.

Graham AM, Rasmussen JM, Rudolph MD, Heim CM, Gilmore JH, Styner M, Potkin SG, Entringer S, Wadhwa PD, Fair DA *et al*. 2018. Maternal systemic Interleukin-6 during pregnancy is associated with newborn amygdala phenotypes and subsequent behavior at 2 years of age. *Biol Psychiatry*. 83(2):109–119. doi: 10.1016/j.biopsych.2017.05.027.

Gross J, Ligges U. 2015. Nortest: tests for normality. R package version 1.0–4.

Hajek T, Kopecek M, Kozeny J, Gunde E, Alda M, Höschl C. 2009. Amygdala volumes in mood disorders — meta-analysis of magnetic resonance volumetry studies. *J Affect Disord*. 115(3):395–410. doi: 10.1016/j.jad.2008.10.007.

Halldorsdottir T, Piechaczek C, Soares de Matos AP, Czamara D, Pehl V, Wagenbuechler P, Feldmann L, Quickenstedt-Reinhardt P, Allgaier A-K, Freisleder FJ *et al*. 2019. Polygenic risk: predicting depression outcomes in clinical and epidemiological cohorts of youths. *Am J Psychiatry*. 176(8):615–625. doi: 10.1176/appi.ajp.2019.18091014.

Hamilton JP, Gottlib IH. 2008. Neural substrates of increased memory sensitivity for negative stimuli in major depression. *Biol Psychiatry*. 63(12):1155–1162. doi: 10.1016/j.biopsych.2007.12.015.

Hashempour N, Tuulari JJ, Merisaari H, Lidauer K, Luukkonen I, Saunavaara J, Parkkola R, Lähdesmäki T, Lehtola SJ, Keskinen M et al.. 2019. A novel approach for manual segmentation of the amygdala and hippocampus in neonate MRI. *Frontiers in Neuroscience*. 13(1025):1–15. doi: 10.3389/fnins.2019.01025.

Insausti R, Amaral DG. 2012. Hippocampal Formation. In: Paxinos G, Mai JK, editors. The Human Nervous System. 3rd ed. San Diego: Academic Press, pp. 896–942.

Karlsson L, Tolvanen M, Scheinin NM, Uusitupa H-M, Korja R, Ekholm E, Tuulari JJ, Pajulo M, Huotilainen M, Paunio T *et al*. 2018. Cohort profile: the FinnBrain birth cohort study (FinnBrain). *Int J Epidemiol*. 47(1):15–16. doi: 10.1093/ije/dyx173.

Kim DR, Bale TL, Epperson CN. 2015. Prenatal programming of mental illness: current understanding of relationship and mechanisms. *Curr Psychiatry Rep*. 17(2):5. doi: 10.1007/s11920-014-0546-9.

Kouwenhoven E, Giezen M, Struikmans H. 2009. Measuring the similarity of target volume delineations independent of the number of observers. *Phys Med Biol*. 54(9):2863–2873. doi: 10.1088/0031-9155/54/9/018.

Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Murray CJL. 2018. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*. 392(10159):1859–1922. doi: 10.1016/S0140-6736(18)32335-3.

Lebel C, Walton M, Letourneau N, Giesbrecht GF, Kaplan BJ, Dewey D. 2016. Pre- and post-partum maternal depressive symptoms are related to children's brain structure in preschool. *Biological Psychiatry*. 80(11):859–868. doi: 10.1016/j.biopsych.2015.12.004.

Lehtola SJ, Tuulari JJ, Scheinin NM, Karlsson L, Parkkola R, Merisaari H, Lewis J, Fonov V, Collins D, Evans A *et al*. n.d. Newborn amygdalar and hippocampal volumes are associated with maternal prenatal psychological distress in a sex-dependent way. (unpublished data).

Lehtola SJ, Tuulari JJ, Karlsson L, Parkkola R, Merisaari H, Saunavaara J, Lähdesmäki T, Scheinin NM, Karlsson H. 2019. Associations of age and sex with brain volumes and asymmetry in 2–5-week-old infants. *Brain Struct Funct*. 224(1):501–513. doi: 10.1007/s00429-018-1787-x.

Lewis JD, Fonov VS, Collins DL, Evans AC, Tohka J. 2019. Cortical and subcortical T1 white/gray contrast, chronological age, and cognitive performance. *Neuroimage*. 196:276–288. doi: 10.1016/J.NEUROIMAGE.2019.04.022.

Lewontin R, Hartl D. 1991. Population genetics in forensic DNA typing. *Science*. 254(5039):1745–1750. doi: 10.1126/science.1845040.

Lieberman MD, Cunningham WA. 2009. Type I and type II error concerns in fMRI research: re-balancing the scale. *Soc Cogn Affect Neurosci*. 4(4):423–428. doi: 10.1093/scan/nsp052.

Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. 2012. The brain basis of emotion: a meta-analytic review. *Behav Brain Sci*. 35(3):121–202. doi: 10.1017/S0140525X11000446.

Loh PR, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, Schoenherr S, Forer L, McCarthy S, Abecasis GR *et al*. 2016. Reference-based phasing using the haplotype reference consortium panel. *Nat Genet*. 48(11):1443–1448. doi: 10.1038/ng.3679.

Lubke GH, Hottenga JJ, Walters R, Laurin C, De Geus EJC, Willemsen G, Smit JH, Middeldorp CM, Penninx BWJH, Vink JM *et al*. 2012. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biol Psychiatry*. 72(8):707–709. doi: 10.1016/j.biopsych.2012.03.011.

Martin A, Kanai M, Kamatani Y, Okada Y, Neale B, Daly M. 2018. Current clinical use of polygenic scores will risk exacerbating health disparities. *bioRxiv*. 176(1):441261. doi: 10.1101/441261.

McKinnon MC, Yucel K, Nazarov A, MacQueen GM. 2009. A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. *J Psychiatry Neurosci*. 34(1):41–54.

Merke DP, Giedd JN, Keil MF, Mehlinger SL, Wiggs EA, Holzer S, Rawson E, Vaituzis AC, Stratakis CA, Chrousos GP. 2005. Children experience cognitive decline despite reversal of brain atrophy one year after resolution of Cushing syndrome. *J Clin Endocrinol Metab*. 90(5):2531–2536. doi: 10.1210/jc.2004-2488.

Meyer-Lindenberg A, Weinberger DR. 2006. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci*. 7(10):818–827. doi: 10.1038/nrn1993.

Naaijen J, Mulder LM, Ilbegi S, de Bruijn S, Kleine-Deters R, Dietrich A, Hoekstra PJ, Marsman J-B, Aggensteiner PM, Holz NE *et al*. 2018. Reactive/proactive aggression specific cortical and subcortical alterations in children and adolescents with disruptive behavior. bioRxiv. doi: 10.1101/490086. [accessed 2019 Feb 8. https://www.biorxiv.org/content/10.1101/490086v1.abstract.

Nederhof E, Schmidt MV. 2012. Mismatch or cumulative stress: toward an integrated hypothesis of programming effects. *Physiology and Behavior*. 106(5):691–700. doi: 10.1016/j.physbeh.2011.12.008.

O'Donnell KJ, Meaney MJ. 2017. Fetal origins of mental health: the developmental origins of health and

disease hypothesis. *Am J Psychiatry*. 174(4):319–328. doi: 10.1176/appi.ajp.2016.16020138.

Pagliaccio D, Luby J, Gaffrey M, Belden A, Botteron K, Gotlib IH, Barch DM. 2012. Anomalous functional brain activation following negative mood induction in children with pre-school onset major depression. *Dev Cogn Neurosci*. 2(2):256–267. doi: 10.1016/j.dcn.2011.11.008.

Pagliaccio D, Luby JL, Luking KR, Belden AC, Barch DM. 2014. Brain–behavior relationships in the experience and regulation of negative emotion in healthy children: implications for risk for childhood depression. *Dev Psychopathol*. 26:1289–1303. doi: 10.1017/S0954579414001035.

Peluso MAM, Glahn DC, Matsuo K, Monkul ES, Najt P, Zamarripa F, Soares JC. 2009. Amygdala hyperactivation in untreated depressed individuals. *Psychiatry Research*. 173(2):158–161. doi: 10.1016/j.pscychresns.2009.03.006.

Peyrot WJ, Milaneschi Y, Abdellaoui A, Sullivan PF, Hottenga JJ, Boomsma DI, Penninx BWJH. 2014. Effect of polygenic risk scores on depression in childhood trauma. *Br J Psychiatry*. 205(2):113–119. doi: 10.1192/bjp.bp.113.143081.

Porcelli S, Fabbri C, Serretti A. 2012. Meta-analysis of serotonin transporter gene promoter polymorphism (5-HTTLPR) association with antidepressant efficacy. *European Neuropsychopharmacology*. 22(4):239–258. doi: 10.1016/j.euroneuro.2011.10.003.

Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, Reich D. 2007. A genomewide admixture map for latino populations. *American Journal of Human Genetics*. 80(6):1024–1036. doi: 10.1086/518313.

van der Plas EAA, Boes AD, Wemmie JA, Tranel D, Nopoulos P. 2010. Amygdala volume correlates positively with fearfulness in normal healthy girls. *Soc Cogn Affect Neurosci*. 5(4):424–431. doi: 10.1093/scan/nsq009.

Qiu A, Shen M, Buss C, Chong YS, Kwek K, Saw SM, Gluckman PD, Wadhwa PD, Entringer S, Styner M *et al*. 2017. Effects of antenatal maternal depressive symptoms and socio-economic status on neonatal brain development are modulated by genetic risk. *Cereb Cortex*. 27(5):3080–3092. doi: 10.1093/cercor/bhx065.

R Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Core Team.

Rao U, Chen LA, Bidesi AS, Shad MU, Thomas MA, Hammen CL. 2010. Hippocampal changes associated with early-life adversity and vulnerability to depression. *Biol Psychiatry*. 67(4):357–364. doi: 10.1016/j.biopsych.2009.10.017.

Revelle W. 2018. *Psych: procedures for personality and psychological research*. Software. Evanston, Illinois, USA: Northwestern University. https://CRAN.R-project.org/package=psych

Rice F, Harold GT, Boivin J, Van Den Bree M, Hay DF, Thapar A. 2010. The links between prenatal stress and offspring development and psychopathology: disentangling environmental and inherited influences. *Psychol Med*. 40(2):335–345. doi: 10.1017/S0033291709005911.

Rifkin-Graboi A, Bai J, Chen H, Hameed WB, Sim LW, Tint MT, Leutscher-Broekman B, Chong Y-S, Gluckman PD, Fortier MV *et al*. 2013. Prenatal maternal depression associates with microstructure of right amygdala in neonates at birth. *Biol Psychiatry*. 74(11):837–844. doi: 10.1016/j.biopsych.2013.06.019.

Ripke S, Daly J, Lewis CM, Lin D, Wray NR, Neale B, William A. 2013. A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*. 18(4):497–511. doi: 10.1038/mp.2012.21.

Roberson-Nay R, McClure EB, Monk CS, Nelson EE, Guyer AE, Fromm SJ, Charney DS, Leibenluft E, Blair J, Ernst M *et al*. 2006. Increased amygdala activity during successful memory encoding in adolescent major depressive disorder: an fMRI study. *Biol Psychiatry*. 60(9):966–973. doi: 10.1016/j.biopsych.2006.02.018.

Sandman CA, Buss C, Head K, Davis EP. 2015. Fetal exposure to maternal depressive symptoms is associated with cortical thickness in late childhood. *Biological Psychiatry*. 77(4):324–334. doi: 10.1016/j.biopsych.2014.06.025.

Siegle GJ, Steinhauer SR, Thase ME, Stenger VA, Carter CS. 2002. Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biological Psychiatry*. 51(9):693–707. doi: 10.1016/S0006-3223(02)01314-8.

Soe NN, Wen DJ, Poh JS, Chong Y-S, Broekman BF, Chen H, Shek LP, Tan KH, Gluckman PD, Fortier MV *et al*. 2018. Perinatal maternal depressive symptoms alter amygdala functional connectivity in girls. *Hum Brain Mapp*. 39(2):680–690. doi: 10.1002/hbm.23873.

Stekhoven DJ, Bühlmann P. 2012. Missforest-non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 28(1):112–118. doi: 10.1093/bioinformatics/btr597.

Sullivan P, Neale MC, Kendler KS. 2000. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry*. 157(10):1552–1562. doi: 10.1176/appi.ajp.157.10.1552.

Suzuki H, Botteron KN, Luby JL, Belden AC, Gaffrey MS, Babb CM, Nishino T, Miller MI, Ratnanather JT, Barch DM. 2013. Structural-functional correlations between hippocampal volume and cortico-limbic emotional responses in depressed children. *Cogn Affect Behav Neurosci*. 13(1):135–151. doi: 10.3758/s13415-012-0121-y.

Taylor MJ, Martin J, Lu Y, Brikell I, Lundström S, Larsson H, Lichtenstein P. 2019. Association of genetic risk factors for psychiatric disorders and traits of these disorders in a Swedish population twin sample. *JAMA Psychiat*. 76(3):280–289. doi: 10.1001/jamapsychiatry.2018.3652.

Thomas KM, Drevets WC, Dahl RE, Ryan ND, Birmaher B, Eccard CH, Casey BJ. 2001. Amygdala response to fearful faces in anxious and depressed children. *Archives of General Psychiatry*. 58(11):1057–1063. doi: 10.1001/archpsyc.58.11.1057.

Tottenham N, Hare TA, Quinn BT, McCarry TW, Nurse M, Gilhooly T, Millner A, Galvan A, Davidson MC, Eigsti IM *et al*. 2010. Prolonged institutional rearing is associated with atypically large amygdala volume and difficulties in emotion regulation. *Dev Sci*. 13(1):46–61. doi: 10.1111/j.1467-7687.2009.00852.x.

Townsend JD, Eberhart NK, Bookheimer SY, Eisenberger NI, Foland-Ross LC, Cook IA, Sugar CA, Altshuler LL. 2010. fMRI activation in the amygdala and the orbitofrontal cortex in unmedicated subjects with major depressive disorder. *Psychiatry Res*. 183(3):209–217. doi: 10.1016/j.pscychresns.2010.06.001.

Visser M, Müller DMJ, van Duijn RJM, Smits M, Verburg N, Hendriks EJ, Nabuurs RJA, Bot JCJ, Eijgelaar RS, Witte M *et al*. 2019. Inter-rater agreement in glioma segmentations on longitudinal MRI. *NeuroImage Clin*. 22(July 2018):101727. doi: 10.1016/j.nicl.2019.101727.

Vyas A, Mitra R, Shankaranarayana Rao BS, Chattarji S. 2002. Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *The Journal of Neuroscience : The Official Journal of the Society for*

Neuroscience. 22(15):6810–6818. doi: 10.1523/JNEUROSCI.22-15-06810.2002.

Wang C, Shen M, Guillaume B, Chong Y, Chen H, Fortier MV, Meaney MJ, Qiu A. 2018. FKBP5 moderates the association between antenatal maternal depressive symptoms and neonatal brain morphology. *Neuropsychopharmacology*. 43(3):564–570. doi: 10.1038/npp.2017.232.

Ward JH. 1963. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc*. 58(301):236–244. doi: 10.1080/01621459.1963.10500845.

Weier K, Fonov V, Lavoie K, Doyon J, Louis CD. 2014. Rapid automatic segmentation of the human cerebellum and its lobules (RASCAL)-implementation and application of the patch-based label-fusion technique with a template library to segment the human cerebellum. *Hum Brain Mapp*. 35(10):5026–5039. doi: 10.1002/hbm.22529.

Wen DJ, Poh JS, Ni SN, Chong Y-S, Chen H, Kwek K, Shek LP, Gluckman PD, Fortier MV, Meaney MJ *et al*. 2017. Influences of prenatal and postnatal maternal depression on amygdala volume and microstructure in young children. *Transl Psychiatry*. 7(4):e1103. doi: 10.1038/tp.2017.74.

Wickham H. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer Verlag. doi: 10.1007/978-0-387-98141-3.

Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, Olesen J, Allgulander C, Alonso J, Faravelli C *et al*. 2011. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 21(9):655–679. doi: 10.1016/j.euroneuro.2011.07.018.

Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF *et al*. 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 50(5):668–681. doi: 10.1038/s41588-018-0090-3.

Yilmazer-Hanke DM. 2012. Amygdala. In: Mai JK, Paxinos G, editors. *The Human Nervous System*. 3rd ed. San Diego: Academic Press. pp. 759–834.