

ARTICLE OPEN



Association between polygenic risk for Major Depression and brain structure in a mega-analysis of 50,975 participants across 11 studies

Xueyi Shen ^{1✉}, Yara J. Toenders ², Laura K. M. Han ^{3,4,5}, Antoine Weihs ^{6,7}, Nina Alexander⁸, Till F. M. Andlauer ⁹, Katharina Brosch ⁸, Andreas J. Forstner ^{10,11,12}, Dominik Grotegerd¹³, Tim Hahn ¹³, Marco Hermesdorf¹⁴, Norbert Hosten¹⁵, Hamidreza Jamalabadi ⁸, Susanne Meinert ^{13,16}, Yuri Milaneschi ^{5,17,18}, Philipp G. Sämann ¹⁹, Frederike Stein⁸, Aleks Stolicyn ¹, Lea Teutenberg⁸, Gladi Thng¹, Mark J. Adams ¹, Florian Thomas-Odenthal ⁸, Paula Usemann⁸, Uwe Völker²⁰, Katharina Wittfeld ⁶, Marisol Herrera-Rivero ¹⁴, Yunxuan Jiang²¹, Chao Tian²¹, on behalf of the 23andMe Research Team*, Nynke A. Groenewold ²², Sheri-Michelle Koopowitz²², Lachlan T. Strike²³, Udo Dannlowski ¹³, Andreas Jansen⁸, Tilo Kircher ⁸, Igor Nenadić ⁸, Kang Sim^{24,25,26}, Benjamin Straube ⁸, Henry Völzke²⁷, Dan J. Stein ^{22,28}, Sarah E. Medland ²³, Klaus Berger¹⁴, Hans J. Grabe⁶, Axel Krug ²⁹, Katie L. McMahon ³⁰, Greig de Zubicaray ³¹, Elena Pozzi^{3,4}, Dick J. Veltman³², Sophia I. Thomopoulos³³, Neda Jahanshad³³, Paul M. Thompson³³, Lianne Schmaal ^{3,4}, Andrew M. McIntosh ^{1,34}, Heather C. Whalley ^{1,34} and for the ENIGMA MDD working group*

© The Author(s) 2025

Major Depression (MD) is a prevalent, disabling and life-limiting condition. The neurobiological associations of genetic risk for MD remain under-explored in large samples, with no comprehensive mega-analysis conducted to date. Our study analysed data from 11 separate studies, encompassing 50,975 participants from the ENIGMA Major Depressive Disorder Working Group. We developed highly consistent genetic and neuroimaging protocols and applied these throughout all participating studies, together with rigorous genetic methods to remove overlap between the polygenic risk scores (PRS) training and testing samples. Elevated PRS for MD correlated with lower intracranial volume and lower global measure of cortical surface area ($\beta_{ICV} = -0.017$, $p_{ICV} = 1.97 \times 10^{-6}$; $\beta_{Surf} = -0.013$, $p_{Surf} = 4.5 \times 10^{-4}$; $pFDR < 3.62 \times 10^{-4}$). The most significant cortical association was observed in the surface area of the frontal lobe ($\beta = -0.011$, $p = 2.85 \times 10^{-6}$, $pFDR = 1.42 \times 10^{-5}$), particularly in the left medial orbito-frontal gyrus ($\beta = -0.021$, $p = 9.48 \times 10^{-8}$, $pFDR = 1.25 \times 10^{-5}$). In subcortical regions, lower volumes of the thalamus, hippocampus, and pallidum correlated with higher PRS of MD (β ranged from -0.011 to -0.015 , p ranged from 0.002 – 1.73×10^{-5} , $pFDR < 0.006$). In a subsample of young individuals only (<25 years old, $N = 5570$), although there were no FDR-significant findings, directions of effects were highly consistent between the analyses of cortical surface areas in youth and the full sample (71.2% in the same direction, exact binomial test p -value = 7.56×10^{-4}). Subsequent Mendelian randomisation analysis revealed potentially causal effects of smaller left hippocampal volume on higher liability for MD (Inverse variance weighted analysis $\beta = -0.064$, $p = 8.04 \times 10^{-3}$, $pFDR = 0.04$). Our findings represent an example of how extensive international collaborations can significantly advance our neurogenetic understanding of MD and give insights to avenues for early interventions in those at high risk for developing MD.

Molecular Psychiatry (2026) 31:611–621; <https://doi.org/10.1038/s41380-025-03136-4>

INTRODUCTION

Major Depression (MD) is a leading contributor to global health burden, costing \$0.7 trillion per year globally [1]. MD is a heritable condition with a twin-based heritability of 37% [2] and a SNP-estimated heritability of 6% [3]. Recent large genome-wide association studies (GWAS) of MD, a broader trait that is highly genetically correlated with clinically defined Major Depressive Disorder, have enabled the creation of polygenic risk scores (PRS) that can objectively estimate liability for MD [3]. Further, PRS can be used to facilitate association analyses in a wide range of health outcomes and biomarkers, and thereby help identify underlying disease mechanisms [4].

Previous large-scale studies have revealed cortical and sub-cortical structural brain abnormalities associated with MD diagnosis [5, 6], including case-control differences in hippocampal volumes [7]. Another meta-analysis revealed that MD was associated with lower cortical thickness in orbitofrontal cortex, cingulate, insula, and the temporal lobes [5]. In GWAS of MD, strong evidence has emerged for enrichment of brain-expressed genes and those involved in synaptic pathways [3]. These findings suggest that the structural brain differences and MD may have shared genetic architecture [8], although this has not been explored on a large scale.

A full list of author affiliations appears at the end of the paper.

Received: 20 June 2024 Revised: 5 May 2025 Accepted: 28 July 2025
Published online: 19 August 2025

Although neuroanatomical associations with genetic risk of MD have been implicated in previous studies, there have been limited large-scale studies examining the association between genetic risk and neuroimaging measures in MD, particularly across multiple cohorts [9]. Previous single-cohort studies often lacked sufficient statistical power to detect reliable associations with small effect sizes [10], and therefore tended to produce inconclusive findings in varying brain regions and inconsistent directions of effect [11]. Differences in pre-processing of imaging data between studies also introduce heterogeneity and may lead to lower reproducibility [12], especially in small brain regions [13]. Additionally, there has been no individual-level mega-analysis looking at the genetic risk of MD and structural brain variations [11]. The lack of large-scale neuroimaging and genetic datasets and the difficulty in harmonising analysis protocols have been major obstacles contributing to the lack of well-powered studies of this type [14, 15], and explains why this approach has rarely been implemented. Finally, previous studies mainly focused on mid-to late-life adults, with limited analyses of younger individuals [16]. This leads to difficulty in generalising findings across the lifespan [16]. Thus, a large-scale, multi-cohort study with representation across a wide age range is a crucial step to deepen our understanding of the neurogenetic basis of MD, and to potentially allow further identification of causal neural biomarkers.

Here, we conducted a unique, individual-level mega-analysis ($N = 50,975$) of participants from 11 studies to examine associations between PRS for MD, cortical thickness and surface area, and subcortical volumes. Standardised imaging and genetic protocols were used to harmonise genetic and neuroimaging data across cohorts. Our study sample included participants from a wide age range (mean age per cohort varied between 9–65 years old), with follow-up analyses that particularly focused on young individuals aged ≤ 25 years ($N = 5570$). For structural brain measures associated with PRS for MD, a follow-up bidirectional Mendelian randomisation analysis was conducted to test if there was any potential causal relationship between structural variation and liability for MD.

METHODS

Participants

A total of 50,975 participants with European ancestry from nine cohorts of the ENIGMA Major Depressive Disorder working group [5, 7], UK Biobank [17] and ABCD [18] datasets were included in the analysis. All ENIGMA cohorts and UK Biobank mainly consisted of adult participants, and ABCD was a youth cohort. Demographic information for each individual cohort can be found in Supplementary Table 1. All cohorts obtained approval from their local institutional review boards and ethics committees. Written consent was obtained from all participants and their caregivers if the participants were minors. The study was approved by the NHS Tayside Research Ethics committee (05/s1401/89). Ethics approval for each individual cohort can be found in their protocol papers and in Supplementary Materials (Supplementary Methods section).

Summary statistics for MD GWAS

This study used Howard et al. (2019) MD GWAS summary statistics [3], excluding 23andMe and all overlapping individuals with the ENIGMA/UK Biobank imaging/ABCD datasets. Although MD is a broad trait that include Major Depressive Disorder, it showed high genetic correlation with clinically ascertained Major Depressive Disorder ($r_G = 0.86$) [3]. Specifically, potentially overlapping individuals between the GWAS and testing samples were identified and then removed using the Psychiatric Genetic Consortium (PGC) CheckSums algorithm [3]. The CheckSums algorithm is an established algorithm that can help identify overlapping samples anonymously. An analysis proposal regarding this study was approved by PGC to allow removing potential overlapping samples. The CheckSums scripts were sent to all the ENIGMA cohorts and the outputs were then shared with the PGC GWAS analyst. Therefore, no ENIGMA study had access to the CheckSums data of PGC or of any other ENIGMA study. Participants that were both part of the GWAS and the ENIGMA Major Depressive Disorder consortium were removed from the discovery GWAS, and an updated analysis of the summary statistics was conducted by a PGC analyst. ABCD study was not included in the MD GWAS as the age group was significantly younger than all participants in the GWAS and therefore no CheckSums analysis was needed. The final MD GWAS summary statistics covered 727,742 individuals. These MD GWAS summary statistics were subsequently used to calculate MD PRS in independent participants.

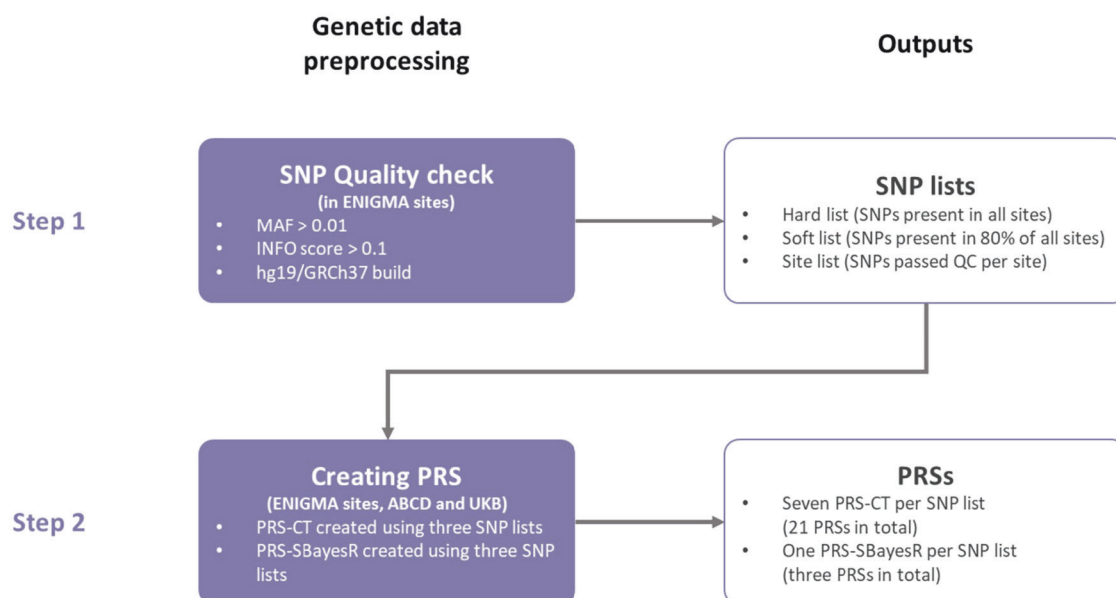


Fig. 1 Flowchart of genetic data processing. PRS processing was performed in two steps. The first step contained standardised quality check for genetic data and SNP selection. Specific criteria are summarised in the figure and detailed in the Methods. The outputs from the first step were then used for the second step of generating PRS for each individual cohort. All PRS created are summarised in the figure and detailed in the Methods. QC quality check, MAF minor allele frequency, PRS polygenic risk score, PRS-CT PRS created using clumping and thresholding method, PRS-SBayesR PRS created using SBayesR.

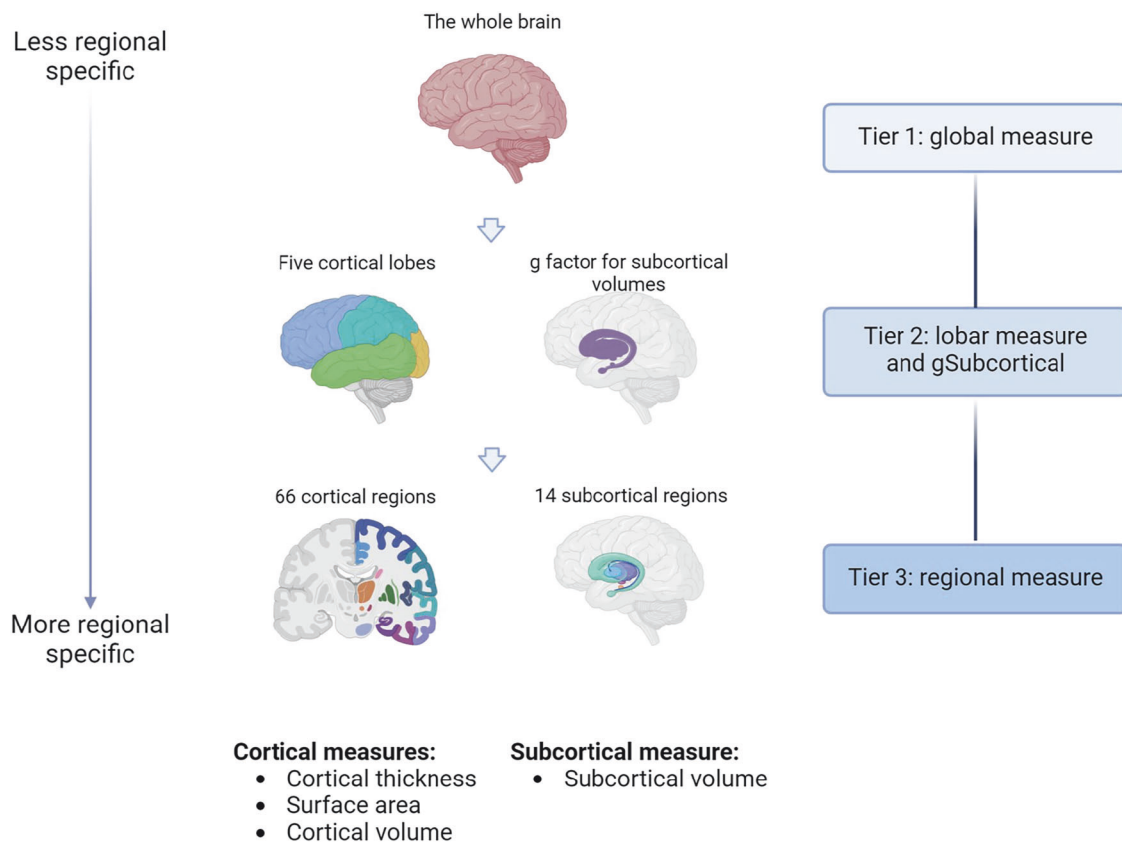


Fig. 2 Neuroimaging measures tested. Associations between MD PRS and brain structural measures were tested hierarchically in three tiers. Tier 1 includes global measures, tier 2 consists of lobar structural measures (cortical thickness and surface area) or the entire subcortical structure, and finally, tier 3 consists of regional measures (cortical thickness, surface area and subcortical volume) of each individual region.

Genetic data processing

Quality check (QC) and pre-processing were conducted locally by each individual cohort providing access to linked genetic and imaging data, with only anonymised individual-level PRS data being shared for analyses (Fig. 1). Hard-call, imputed genetic data were used for generating PRS. Selection of SNPs included in creating PRS was performed in a two-stage manner. First, all ENIGMA cohorts shared a list of SNPs in the hg19/GRCh37 build that passed the QC criteria of minor allele frequency (MAF) > 0.01 and INFO score > 0.1. Based on these SNP lists, we used three lists of SNPs for each individual cohort to create PRS: (1) a hard list: SNPs present in all ENIGMA cohorts ($N_{\text{SNP}} = 3,176,977$), (2) a soft list: SNPs present in more than 80% of the cohorts in ENIGMA ($N_{\text{SNP}} = 6,306,997$) and (3) a cohort list: all SNPs that passed QC for the specific individual cohort (N_{SNP} varies per cohort). These three lists were also applied to select SNPs in the UK Biobank and ABCD. For cohort lists in the UK Biobank and ABCD, a lower threshold of minor allele frequency (MAF > 0.001) was used due to the large sample sizes of these cohorts.

Calculation of PRS for MD

We created PRS for MD using two methods. The first method we used was the clumping and thresholding method (PRS-CT) [19]. We created seven PRS-CT using PRSice 2.0 [19] based on seven p-value thresholds ($p_T = 5 \times 10^{-8}, 1 \times 10^{-4}, 1 \times 10^{-3}, 0.01, 0.1, 0.5$ and 1). Clumping was conducted using a window of 500 kb with a $r^2 < 0.1$. For smaller studies with $N < 1000$, genotype data of central European samples (CEU) from the 1000 Genomes Project was used as reference panel for clumping [20]. For larger studies of $N > 1000$, their own imputed genetic data was used as reference data for clumping.

We also created a second set of PRS using a Bayesian method (PRS-SBayesR) [21]. The Bayesian method of creating PRS has become increasingly used due to its better predictive power. MD summary statistics were processed using 'SBayesR' [21], and the shrunk sparse LD matrix that covers 2.8 million common SNPs was used as the LD reference panel [21]. Markov chain Monte Carlo (MCMC) chain length was set as 21,000. The output summary statistics of SBayesR were then used for creating PRS. A total of

86,787 SNPs with non-zero effect sizes were included in the SBayesR summary statistics. The summary statistics were then used in PRSice 2.0 to create the PRS, and no clumping was applied. As a result, six sets of PRS were generated per cohort (three SNP lists \times two PRS methods). Scripts are available at the URL: https://github.com/xshen796/ENIGMA_mdd_prs/blob/main/script/PREP_PRS/Calculate_PRS.md.

Neuroimaging measures

T1-weighted images were obtained and pre-processed locally for each individual cohort [5, 7]. Anonymised individual-level FreeSurfer outputs were shared for analysis. The T1-weighted images were preprocessed, quality-checked and parcellated with FreeSurfer version 5.0 or 5.3 using the ENIGMA3 – GWAS Meta Analysis of Cortical Thickness and Surface Area protocol. Detailed description can be found in the protocol paper [22] and in the URL: <https://enigma.ini.usc.edu/protocols/imaging-protocols/>.

The neuroimaging measures were generated in a three-tier hierarchical order [23] (see Fig. 2): (1) global measures, (2) lobar measures and g factor of subcortical volumes (gSubcor) and (3) regional measures. The global measures include intracranial volume (ICV), global average cortical thickness and total brain surface area. The lobar measures included cortical thickness and surface area estimates of five lobes: temporal, parietal, occipital, frontal and cingulate lobes. Cortical thickness and surface area for each lobe was calculated by extracting the mean of thickness and the sum of surface area for all regions that lie within the lobe. For each participant, lobar measures which had more than two regions with missing values were set as 'NA' and therefore removed from further analysis. Definitions of lobar regions can be found elsewhere [23] and in the Supplementary Materials. The gSubcor [6] measure represents the score on the first unrotated principal component (PC), derived with PCA conducted on all subcortical volumes. The first unrotated PC explains 64.9% of the total variance of all 14 subcortical volumes (Supplementary Fig. 1). Finally, regional measures included cortical thickness and surface area estimates of 33 bilateral cortical regions defined by the 'Desikan-Killiany' cortical atlas [5, 24] (66 measures per participant) as well as subcortical volumes for 7 regions defined by the ASEG subcortical atlas (14

measures per participant) [7, 25]. Lateral ventricles and the temporal pole were not included in the analyses, as these measures were not available in large cohorts (e.g. UKB, see Supplementary Materials).

MD PRS association analysis

Validation of MD PRS prediction. To validate the MD PRS, out-of-sample prediction of all the MD PRS (eight PRS in total, including seven PRS-CT and one PRS-SBayesR) was performed using a mixed-effect linear regression model [26]. Before entering the linear regression model, each MD PRS was residualised against covariates to account for genetic relatedness and population stratification. Different strategies were applied separately for ENIGMA cohorts, the ABCD study and UK Biobank. For ENIGMA cohorts, MD PRS were regressed against genomic relationship matrices. Genomic relationship matrices were derived using the imputed genetic data that passed quality check using GCTA [27], and they were fit as the random factor. The residualised scores were generated per cohort using the 'lme4' function from the 'lme4' R package [28]. For the ABCD study, MD PRS were regressed against genetic principal components, batch, family ID, assessment site and an additional term of 'family ID|assessment site' to reflect the nested structure, as per recommendation by the ABCD team [29]. A mixed-effect, linear model was used for the ABCD study, using the 'lmer' function from the 'glmer' R package. Family ID, assessment site and the nested term were included as random factors, and genetic principal components were set as fixed factors. For UK Biobank, because its genomic relationship matrix was particularly computationally demanding to handle, we regressed the MD PRS against a kinship relationship matrix using a similar approach that was implemented in our previous work [30]. Kinship relationships up to second-degree relatives were derived using the 'King' software [31]. The first ten 10 genetic principal components, genotyping array, scanner site and head positions in the scanner on the x-, y-, z- and table axes [32] were included as additional covariates along with the kinship matrix. Residualised scores for UK Biobank were obtained using the restricted maximum likelihood approach in GCTA [27]. The scores were set to match the standard deviation of the original uncorrected scores so as to keep the scores within the consistent scale with other studies. The above correction of technical confounders and relatedness were performed on each individual site. The corrected data was then merged as one single dataset for further analyses. Correlations between PRS and PRS residualised against covariates for each individual study were reported in Supplementary Table 2 and Supplementary Fig. 2.

Mixed-effect logistic regression was used to test the association between the residualised PRS and MD using the 'glmer' function in the 'lme4' R package [33], adding age and sex as fixed-effect covariates and cohort as a random factor. MD case-control diagnosis was set as the dependent variable (y) and MD PRS were set as the independent variable (x). Standardised log-transformed odds ratios (OR) were reported as effect sizes.

Association between MD PRS and neuroimaging measures. All analyses on associations between MD PRS and brain structural measures were conducted in the full sample first, followed by additional analyses conducted in youth samples (individuals ≤25 years old across all cohorts).

Associations between MD PRS and neuroimaging measures were tested using mixed-effect linear models similar to the model used for MD PRS prediction ('lmer' function in the 'lme4' R package was used). All the covariates and genomic/kinship relationship matrix and technical covariate correction were kept consistent with the MD PRS prediction analyses. Neuroimaging measures were set as the dependent variables and MD PRS as the independent variables. Age and sex were added as fixed-effect covariates and site was added as a random-effect covariate in the linear model.

Associations with MD PRS were tested separately for global, lobar and regional measures (see Fig. 2). PRS which were most strongly associated with global measures amongst all PRS-CT and PRS-SBayesR were further analysed for associations with lobar and regional measures. When regional measures were analysed, an additional model that included ICV as a covariate was tested as secondary analysis to see if the regional effect was over and above global volume. An additional sensitivity analysis was performed to investigate the interaction of MD PRS and MD case-control status in the full sample.

Standardised regression coefficients (β) were reported as effect sizes. P-values were FDR-corrected for lobar and regional measures per type of structural measure (e.g. p-values for cortical thickness were corrected across all 66 regions, and p-values for subcortical volumes were corrected across all

14 regions). We thus refer to this correction as whole-brain FDR correction. We further applied two additional methods of multiple testing correction, which were reported in the Supplementary Materials. The first multiple testing correction method was Bonferroni correction per type of structural measure (i.e. whole-brain Bonferroni correction), and the second method was FDR correction across all measures tested (146 regions in total, i.e. all-measure FDR correction). For global measures, FDR-correction was applied across all PRS and global measures (eight PRS \times three global measures).

To investigate specific associations in younger people, we conducted a separate set of association analyses in young individuals (<25 years old). A total of 5570 young subjects from seven cohorts were included in the analyses (see demographic information in Supplementary Table 1). To formally test the concordance of effect sizes across regions between the associations found in youth and in the full sample, we performed (1) Pearson's correlation analysis on the effect sizes of regional structural measures, and (2) the exact binomial test of directions of effects for regional measures using the 'binom.test' function from the 'stats' R package ('-1' represents a negative effect size and '+1' represents a positive effect size, hypothesised random probability of success set as 0.5).

Mendelian randomisation (MR)

To investigate the potentially causal relationship between brain structure and MD, we conducted bi-directional MR for the brain structural measures that showed significant association with MD PRS. MR utilises genetic variants as instruments for testing causality, and has been particularly useful for identifying causal biomarkers for diseases [34].

MD GWAS. We used the summary statistics from the meta-analysis of the MD GWAS for creating PRS (non-imaging sample of UKB, PGC cohorts) and summary statistics of European subjects from 23andMe, Inc [3]. For the additional summary statistics from 23andMe, participants provided informed consent and volunteered to participate in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of Salus IRB (<https://www.versiticlinicaltrials.org/salusirb>). The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Datasets will be made available at no cost for academic use. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information and to apply to access the data.

GWAS for neuroimaging traits. GWAS summary statistics of neuroimaging traits were extracted from the Smith et al. study of ~33,000 participants with European ancestry. Details of the study can be found elsewhere [3, 4]. The discovery GWAS sample for neuroimaging traits had no overlap with the MD GWAS, which was ensured by using the same protocol for the PRS association analysis. Neuroimaging traits which demonstrated an association with MD PRS and had >10 instrumental genetic variants (IVs) after clumping were included in the MR analysis.

MR data preparation and analysis. We used the 'TwoSampleMR' R package (version 0.5.6) for all MR analyses. Data preparation included standard procedure described elsewhere [4]. In brief, GWAS summary statistics of exposure variables went through QC, and variants were retained if $MAF > 0.01$, $INFO\ score > 0.1$, and Hardy-Weinberg equilibrium $p\text{-value} < 1 \times 10^{-5}$. Exposure summary statistics were then clumped using a $r^2 < 1 \times 10^{-3}$ with a 1 MB window using the default European LD reference panel. Clumping was performed using the 'clump_data' function from the 'TwoSampleMR' package. After quality check and clumping, MD had 122 IVs retained as exposure data, and the number of IVs for neuroimaging traits ranged from 10–24 (median number of IVs = 15).

Three methods were used for bidirectional MR: (1) inverse variance weighted (IVW, the main method), (2) weighted median, and (3) MR Egger (with bootstrapping). For the MR analysis of causal effect of MD, due to the large number of IVs of MD, we used the 'contamination mixture' method as a secondary analysis in addition to IVW. The contamination mixture method has a particular advantage when handling a large number of IVs, and typically results in low Type I error rates when the numbers of invalid IVs are small. Horizontal pleiotropy and heterogeneity of IVs were estimated using Egger intercept and Q statistics, respectively. Those associations were identified as potentially causal association if: (1) IVW/contamination mixture test was FDR-significant, (2) confirmed by nominally significant weighted median results in the same direction of IVW and (3) nominally significant MR Egger analysis if

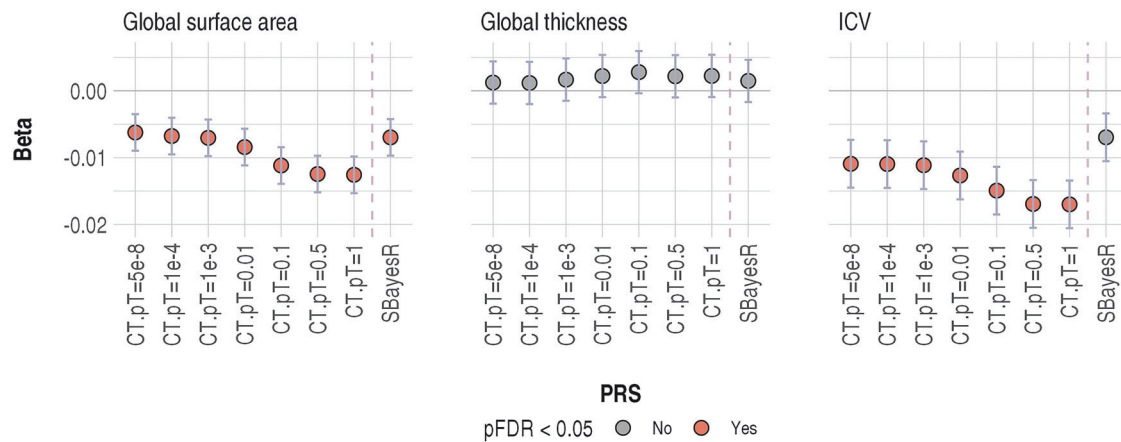


Fig. 3 Association between all MD PRS and global neuroimaging measures. X-axis represents individual PRS. CT clumping and thresholding method, pT p-value threshold used for creating PRS using clumping and threshold methods, and SBayesR PRS created using the SBayesR-processed summary statistics. Y-axis represents β , i.e., standardised regression coefficient.

there was any horizontal pleiotropy indicated by an Egger intercept significantly deviating from zero.

RESULTS

Validation of MD PRS

PRS created using the three sets of SNPs were highly correlated (all $r > 0.99$). Therefore, all the following analyses were conducted using PRS created using only the 'cohort list' variant that was based on SNPs that passed QC criteria for each individual cohort.

For the PRS-CT, those PRS which were created using p-value thresholds higher than 0.001 were associated with MD diagnosis (log-transformed odds ratio ranged from 0.041–0.162, p ranged from 0.002 to 1.5×10^{-34} , Supplementary Table 3). The PRS-SBayesR had stronger association than all the PRS-CT ($\beta = 0.181$, $p = 2.83 \times 10^{-42}$).

MD prediction for each individual study cohort can be found in Supplementary Fig. 3.

Association between MD PRS and global neuroimaging measures

Higher MD PRS were associated with lower ICV for all the PRS-CT (β ranged from -0.011 to -0.017 , p ranged from 0.002 – 1.97×10^{-6} , $pFDR < 0.005$, see Fig. 3, Supplementary Fig. 4 and Supplementary Table 4). The effect size for association between ICV and PRS-SBayesR was in the same direction but did not reach FDR-corrected significance ($\beta = -0.007$, $p = 0.052$, $pFDR = 0.078$). All PRS-CT and PRS-SBayesR were also associated with lower total cortical surface area (β ranged from -0.006 to -0.013 , p ranged from 0.023 – 4.5×10^{-6} , $pFDR < 0.036$). No association was found between any MD PRS and global cortical thickness (β ranged from 0.001 – 0.003 , $p > 0.377$).

Since MD PRS-CT created at the p-value threshold of 1 ($pT = 1$) showed the strongest association with both ICV ($\beta = -0.017$, $p = 1.97 \times 10^{-6}$) and global cortical surface area ($\beta = -0.013$, $p = 4.5 \times 10^{-6}$), analyses for this MD PRS was carried on for all regional measures.

Association between MD PRS-CT ($pT=1$), lobar measures and gSubcortical

Lower surface areas were associated with the MD PRS-CT at $pT = 1$ in all five lobes (β ranged from -0.006 to -0.011 , p ranged from 0.033 – 2.85×10^{-6} , $pFDR < 0.033$, see Supplementary Table 5). The strongest association was found with surface area in the frontal lobe ($\beta = -0.011$, $p = 2.85 \times 10^{-6}$, $pFDR = 1.42 \times 10^{-5}$). Lower gSubcortical was also found associated with the MD PRS ($\beta = -0.011$, $p = 4.41 \times 10^{-4}$).

No lobar measure of cortical thickness was associated with MD PRS-CT at $pT = 1$ (absolute β ranged from 0.002 – 0.005 , $p > 0.149$).

Association between MD PRS-CT ($pT=1$) and regional cortical thickness and surface area

Without correction for ICV, cortical surface areas in 46 regions (out of 66 regions tested) were associated with MD PRS-CT at $pT = 1$ and reached FDR-significance (β ranged from -0.009 to -0.021 , p ranged from 0.031 – 9.48×10^{-8} , $pFDR < 0.044$). Using a more stringent FDR correction across all measures tested, 39 regions remained significant after multiple testing correction (β ranged from -0.01 to -0.021 , p ranged from 0.01 – 9.48×10^{-8} , $pFDR$ across all measures < 0.035). Within the 46 FDR-significant regions, 34 were bilateral regions that had associations in both hemispheres (17 regions of both hemispheres), and 12 regions had associations in one hemisphere only. Strongest associations were found in the medial orbito-frontal gyrus in both hemispheres (left hemisphere: $\beta = -0.021$, $p = 9.48 \times 10^{-8}$, $pFDR = 6.26 \times 10^{-6}$; right hemisphere: $\beta = -0.015$, $p = 5.19 \times 10^{-5}$, $pFDR = 5.83 \times 10^{-4}$) and in bilateral superior frontal gyrus (left hemisphere: $\beta = -0.016$, $p = 2.4 \times 10^{-5}$, $pFDR = 5.29 \times 10^{-4}$; right hemisphere: $\beta = -0.015$, $p = 5.3 \times 10^{-5}$, $pFDR = 5.83 \times 10^{-4}$). See Fig. 4, Supplementary Fig. 5 and Supplementary Tables 6, 7 for regional statistics.

A secondary analysis was carried out on surface areas measures after controlling for ICV (Supplementary Fig. 6). Focusing on the measures that reached FDR-significance in their associations with MD PRS-CT, left medial orbito-frontal gyrus remained significant after controlling for ICV ($\beta = -0.011$, $p = 0.001$, $pFDR = 0.048$). No other regions that were previously associated with MD PRS-CT reached FDR-corrected significance after controlling for ICV ($p > 0.003$, $pFDR > 0.077$).

No association was found between regional cortical thickness and MD PRS-CT ($pT = 1$) (absolute β ranged from 1.96×10^{-4} to 0.008 , $p > 0.023$, $pFDR > 0.7$).

Association between MD PRS ($pT=1$) and subcortical volumes

Without correction for ICV, there were five subcortical regions that showed FDR-significant associations with MD PRS-CT ($pT = 1$), out of the 14 regions tested (β ranged from -0.011 to -0.015 , p ranged from 0.002 – 1.73×10^{-5} , $pFDR < 0.006$, see Supplementary Table 8). All five subcortical regions remained significant using a more stringent FDR-correction across all regional measures ($pFDR$ across all measures < 0.01). Out of all the FDR-significant regions, lower volumes of the thalamus and pallidum were associated with higher MD PRS-CT bilaterally (β ranged from -0.011 to -0.015 , p ranged from 0.001 – 1.73×10^{-5} , $pFDR < 0.004$). Lower volume of the left hippocampus was associated with higher MD PRS

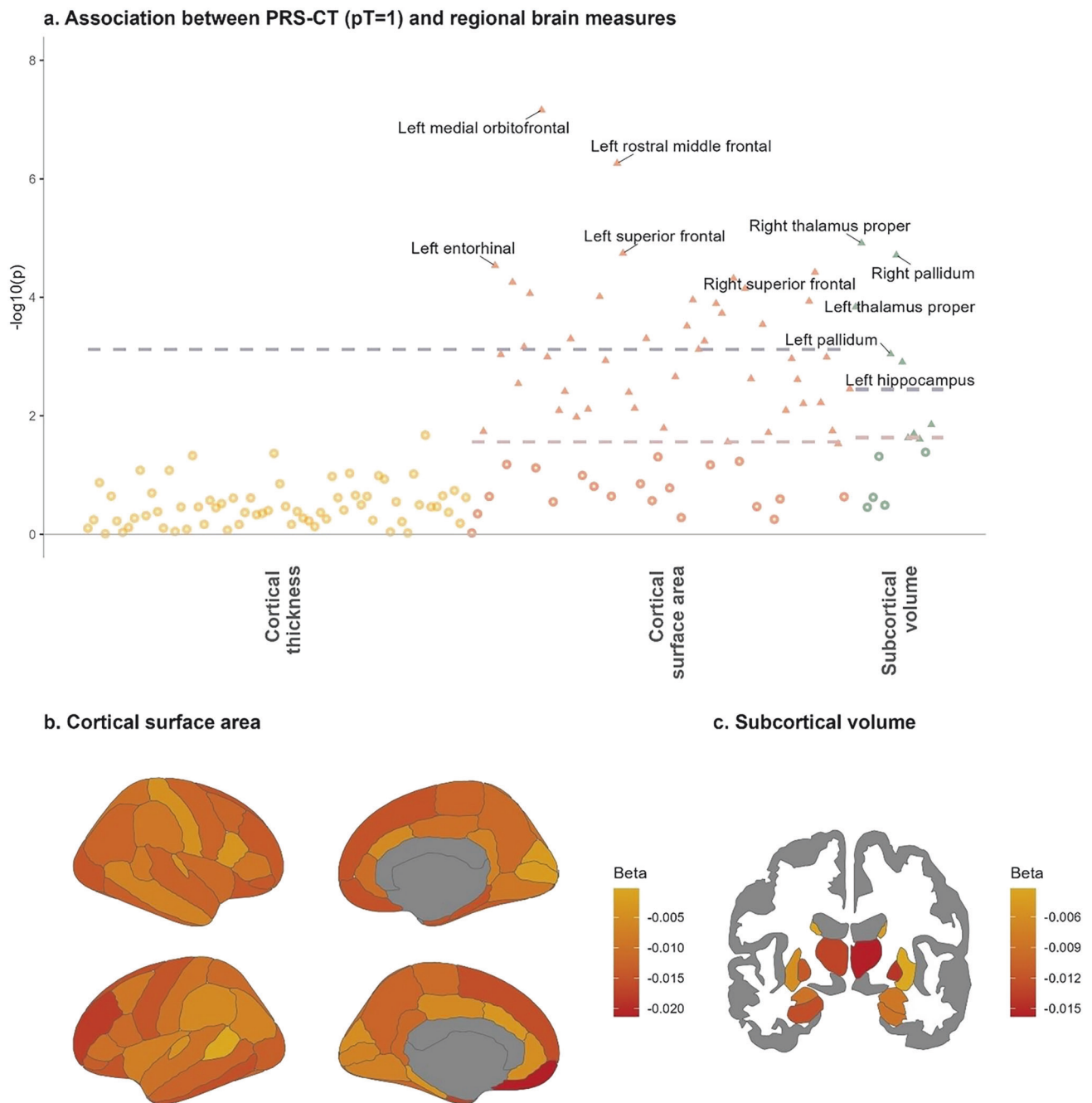


Fig. 4 Association between PRS-CT at pT=1 and brain regional measures. **a** P-value plot for associations between MD PRS and all regional measures. X-axis represents the three categories of brain regional measures. Y-axis represents $-\log_{10}(p)$ -transformed p-values. Each dot represents the result for one brain regional measure. Red and grey dashed lines are the FDR- and Bonferroni-significance threshold, respectively. Associations that are significant after FDR-correction are highlighted as solid dots. Top five FDR-significant associations per category are annotated with labels of regions in the figure. Multiple-comparison correction was applied within each individual category. No associations were found in cortical thickness and therefore only the Bonferroni-significance threshold is shown in the figure. **b**. Regional results for cortical surface areas. Darker colours represent lower beta-values (standardised regression coefficients). **c**. Regional results for subcortical volumes in coronal view.

($\beta = -0.012$, $p = 0.002$, $pFDR = 0.006$). After controlling for ICV in the association model, nominally significant association was found in the right pallidum ($\beta = -0.006$, $p = 0.042$, $pFDR = 0.197$), but no other associations were found ($p > 0.079$).

Interaction between MD PRS and MD case-control status

No interaction between MD PRS and MD case-control status reached FDR-significance in the entire sample for any brain measure ($p > 0.028$, $pFDR > 0.96$).

Analysis specifically in youth (≤ 25 years old)

In addition to the main analyses conducted on the entire sample, we conducted additional analyses on young individuals ≤ 25 years old ($N = 5570$ from seven cohorts, see Supplementary Table 1). For global neuroimaging measures, no association reached FDR-corrected significance (for all cortical measures: absolute β ranged from 2.75×10^{-4} to 0.022 , $p > 0.071$). Similar to the adult sample, effect sizes were larger for ICV and cortical surface area than cortical thickness (see Supplementary Fig. 4 and Supplementary

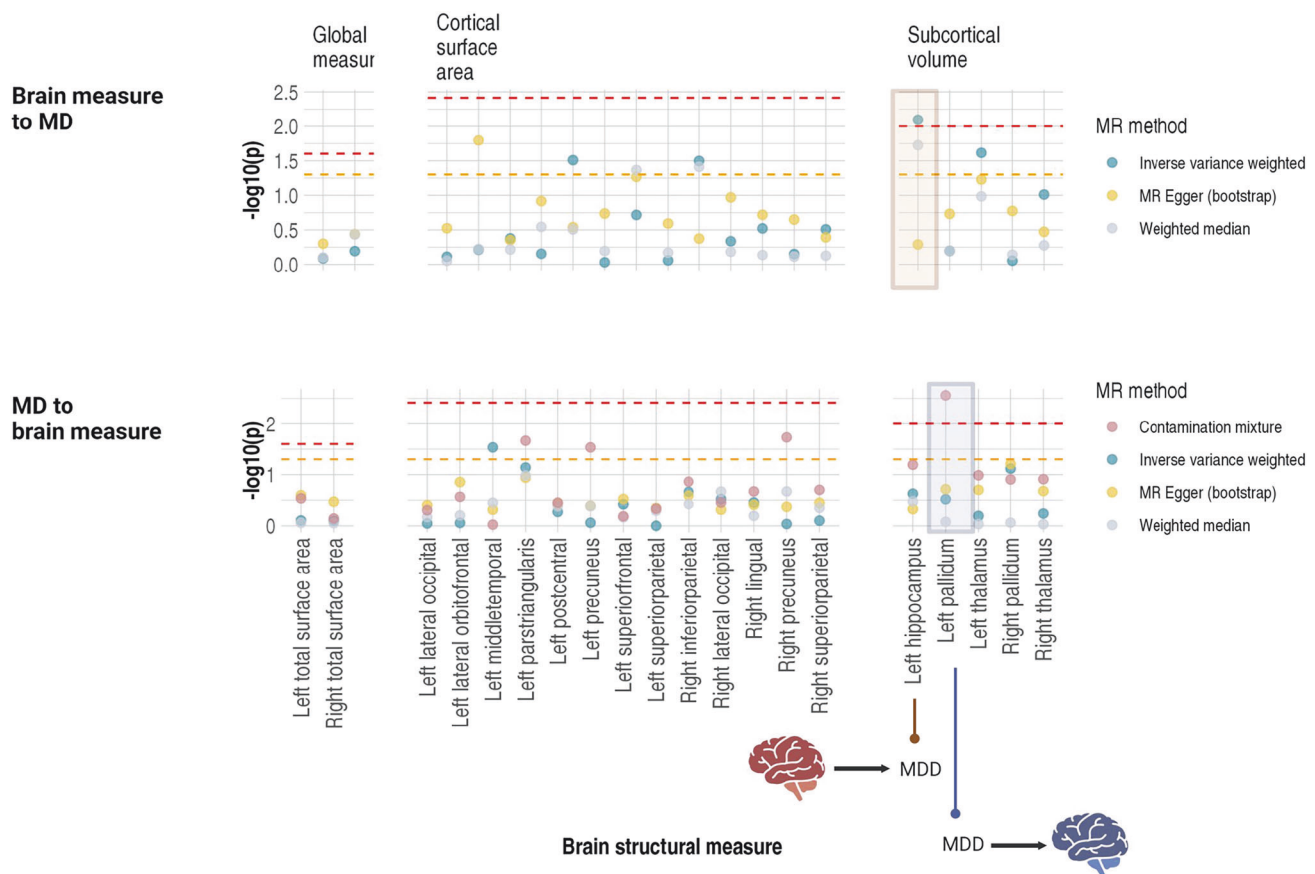


Fig. 5 Results of the bi-directional Mendelian randomisation analysis for the causal relationships between MD and neuroimaging traits. The top panel shows the results of causal effects of brain structural measures to MD, and the bottom panel shows the causal effects in the reverse direction. X-axes represent individual brain structural measures, and the measures were separated by their category (global measure/cortical surface area/subcortical volume). Y-axes represent $-\log_{10}$ -transformed p-values. Neuroimaging traits were highlighted in the plot if there was any potentially causal effect that reached FDR-significance indicated by the inverse variance weighted method.

Table 4). No association was found in lobar/gSubcortical measures (absolute β ranged from 3.55×10^{-4} to 0.013, $p > 0.266$, Supplementary Table 5). However, direction of associations for ICV and global surface area were consistent with associations found in the entire sample (β ranged from -0.007 to -0.022 for ICV, and β ranged from -2.75×10^{-4} to -0.011 for global surface area).

For regional neuroimaging measures, we did not find any associations that reached FDR-corrected significance (absolute β ranged from 5.18×10^{-5} to 0.037, $p > 0.004$, $pFDR > 0.498$, see Supplementary Tables 6–8). The strongest association was found between thickness in the right posterior cingulate gyrus and PRS-CT at $pT = 1$ ($\beta = 0.037$, $p = 0.004$, $pFDR = 0.271$). Although no association reached FDR-significance, effect sizes for regional cortical surface areas were positively correlated between the youth and the full sample ($r = 0.43$, $p = 3.21 \times 10^{-4}$), and the directions of effects were significantly consistent (71.2% associations in the same direction, exact binomial test p -value = 7.56×10^{-4}). This pattern was seen to a lesser degree in cortical thickness, for which effect sizes in youth and the full sample showed a significant but lower correlation ($r = 0.34$, $p = 0.006$), and the exact binomial test suggested that concordance of direction of association was not significantly higher than chance (56.1% associations in the same direction, exact binomial test p -value = 0.389). Similarly, a low concordance rate of direction of association was found for subcortical volumes between the youth and full sample (correlation of effect sizes: $r = 0.752$, $p = 0.002$; 42.9% associations in the same direction, exact binomial test p -value = 0.791).

MR

We identified a potentially causal effect of smaller left hippocampal volume on higher liability for MD (see Fig. 5, Supplementary Fig. 7, Supplementary Tables 9, 10). Left hippocampal volume showed a significant effect on MD in IVW analysis ($\beta = -0.064$, $p = 8.04 \times 10^{-3}$, $pFDR = 0.04$), supported by a nominally significant weighted median result ($\beta = -0.049$, $p = 0.019$). MR Egger analysis was not significant ($\beta = -6.33 \times 10^{-4}$, $p = 0.514$), however, no horizontal pleiotropy was observed (p -value for Egger intercept = 0.346). On the contrary, there was no causal effect identified from MD to left hippocampal volume ($p > 0.064$ for all MR methods, Supplementary Data 1).

No causal relationship in the reverse direction (from MD to neuroimaging traits) was FDR-significant, as assessed with IVW ($p > 0.073$, $pFDR > 0.379$) or weighted median method ($p > 0.104$, $pFDR > 0.684$). The only causal effect found from MD to brain structural trait was in the left pallium using the contamination mixture method ($\beta = -0.179$, $p = 0.003$, $pFDR = 0.014$, No. of valid IVs = 89).

DISCUSSION

We used a large-scale mega-analysis on individual-level data and found that higher PRS for MD were associated with smaller global cortical surface area, lower intracranial volume and lower subcortical volumes, but not cortical thickness. Further analysis of lobar measures indicated that the largest difference associated with MD PRS was in the surface area of the frontal lobe. More

specifically, smaller cortical surface areas in regions including the medial orbito-frontal gyrus, superior frontal gyrus and temporal lobe, were associated with higher MD PRS. In subcortical regions, lower volumes of the thalamus, pallidum and hippocampus were found associated with higher PRS for MD. In young individuals below 25 years old, no associations survived correction for multiple comparison in the relatively small size of the subsample. However, directions of the effects remained similar to those found in the analyses of the entire study sample.

Our study identified brain structural measures that were previously found to have case-control differences in MD, for example, ICV and hippocampal volume [7]. However, the majority of brain structural associations identified in our analysis had not been found consistently in case-control studies. A major difference is that our findings mainly fell within the category of cortical surface areas, whereas cortical thickness differences are more commonly reported in adult MD case-control studies [5]. Our findings of differences in cortical surface areas are, however, supported by findings from studies of genetic correlation [22, 35]. For example, genetic correlations with MD were observed with cortical surface areas, and the associations were stronger than with cortical thickness [22]. Such pattern of stronger association between MD and surface areas than cortical thickness was also observed in youths [36]. The fact that differences were found in surface areas rather than cortical thickness may imply that genetic predisposition to MD may manifest particularly prominently early in life, while abnormality in either cortical thickness or surface area may manifest in different age groups across the life course [37, 38]. Various reasons could contribute to differences in associations of PRS and or diagnosis of MD with the brain, including varying environmental influence across the life course at critical timepoints [35]. Future studies are needed to further explore the gene-by-environment interaction that may contribute to these differences.

One of the strongest regional associations was found in the surface area of the medial orbitofrontal cortex (OFC), and this association remained significant after controlling for ICV. The differences in both structure and functional activity of the OFC in MD has been consistently documented [5, 37]. Structural differences in OFC were previously found to be associated with depressive symptoms in treatment-naïve patients [38] and in pre-school youths [39], suggesting that impairment in OFC may be particularly important in early development of depressive symptoms. Participants with brain lesions in OFC showed lower level of depressive symptoms compared to those with lesions in other brain areas [40]. Moreover, brain stimulation of the OFC area improved mood state in MD patients [41]. These studies suggest a potential causal role of changes in the OFC in MD. Brain activity in OFC is associated with known protective factors for MD, such as reward processing [42, 43] and regulation of stress caused by negative social interactions (e.g. social exclusion) [44, 45]. Anatomically, the lower medial wall of the PFC (medial OFC according to the Desikan–Killiany atlas in FreeSurfer) contains the subgenual anterior cingulate cortex and subcallosal gyrus that serve as homeostasis sensors and control the HPA axis responses through connections to the hypothalamus [46]. Taken together, the functional and structural findings indicate that delayed or abnormal development of OFC may manifest in mood regulation (e.g., reward processing) impairments that are prevalent in MD patients.

In subcortical regions, lower hippocampal volume was previously found associated with lifetime and current MD, and the association was particularly strong in patients with recurrent episodes and in patients with early age of onset < 21 years old [7]. Our findings confirm the crucial role of the hippocampus in MD, as well as its potentially causal relationship with MD [47–49]. Findings from early life adversity and its influence on hippocampal structural integrity also speak to a potentially causal

pathophysiological role of hippocampal deficits for the liability to develop an affective disorder [50]. As such, both genetics and imaging studies provided a potentially causal relationship from hippocampal volume to MD. Further, we report a novel association between PRS for MD and thalamic volume. A previous study looking at genetic correlations found that thalamic volume had overlapping genetic architecture with MD and the genetic correlation was stronger for MD compared with other major psychiatric disorders (e.g. schizophrenia) [35]. Previous large-scale studies found that MD and higher PRS of MD were associated with lower integrity in white matter microstructure of thalamic radiations [4]. In addition to findings related to structural variations, three independent functional MRI studies showed that thalamic activity may be an indicator for treatment outcomes [51–53]. The thalamus is an important hub that connects the fronto-parietal network with the limbic system [54]. Considering its specific anatomical importance in the structural connectome, associations with MD indicate that there is an overall turbulence of thalamic structural/functional network in the disorder [54]. Further studies are needed to directly examine the role of the thalamus and its disruption in the greater neural network.

We are the first to conduct a mega-analysis to investigate associations between genetic risk of MD and brain structure. We took advantage of a large multi-site sample and adopted a hypothesis-free, data-driven approach. This strategy enabled us to reveal novel associations that were not detected in smaller studies, which typically used a region-of-interest approach. A few important sensitivity analyses were conducted to estimate the extent to which heterogeneity of data pre-processing might contribute to between-site differences. In particular, we found that there was minimal difference between the PRS generated using locally quality-checked, site-specific SNP lists and the PRS generated using a single consistent list across sites. Our study is also one of the first to evaluate the influence of different lists of SNPs between cohorts and to implement rigorous examination of sample overlap between training and testing samples in genetic imaging studies. Our findings and the applied methods provide important technical guidance for future neuroimaging-genetic studies and may be used for studying other brain phenotypes, such as those based on functional imaging and diffusion tensor imaging. However, it is important to note that MD is well-known as a highly polygenic trait, and datasets included in the present study were relatively large (all $N > 200$). It is, therefore, unlikely that heterogeneity of SNPs between studies may have had major impact on the results. Caution should be taken especially when considering traits with more sparse genetic architecture influenced by certain genetic variants with relatively stronger effect sizes. And thus inconsistent availability in these genetic variants may pose stronger influence on the heterogeneity in PRS.

The current study has a few limitations. First, we looked only at European samples due to the population constrained GWAS used for creating PRS. It has been found that trans-ancestry PRS tend to have compromised statistical power due to the confounding effect from heterogeneous genetic architectures between ancestry groups [55]. It is imperative that further studies aim to extend the analysis to other ancestry groups when well-powered, non-European GWAS for MD becomes available. More accurate imputation techniques may help maximise the use from samples that have mixed ethnic backgrounds [56]. Second, we utilise MD PRS as key measure since it includes both self-reported and clinically ascertained Major Depressive Disorder. MD is a broader definition of depression, capturing the self-reported core symptoms of MDD, including persistent and severe low mood or despair. MD PRS has demonstrated statistical power to differentiate both self-reported and clinically ascertained cases from controls [3]. Further, multiple versions of FreeSurfer may introduce biases to the results. However, this issue is unlikely to affect larger studies included in our analysis (e.g. UKBB and ABCD). Finally, the

young samples in the present study are relatively small which limits the statistical power to find reliable associations in this age group. Although we included participants from a wide age range, participants in early adulthood were underrepresented [16]. Future studies may benefit from focusing on younger participants and studying disease progression into later life [35].

In conclusion, our analysis identified novel associations between PRS for MD, global cortical surface area and intracranial volume. Genetic risk of MD was, in particular, associated with lower surface area in OFC, and lower hippocampal and thalamic volumes. It was not, however, associated with differences in cortical thickness. For the hippocampus, MR analysis demonstrated a causal effect from hippocampal volume to MD. Previous meta-analysis did not report MD case-control differences in surface areas [5], but our findings are convergent with those of a previous study of genetic architecture of longitudinal brain structural changes [35]. Future studies are necessary to investigate longitudinal changes in the cortical regions identified here, and to define the effects of gene-by-environment interactions on these regions.

DATA AVAILABILITY

GWAS summary statistics for Major Depression are publicly available at the URL: <https://datashare.ed.ac.uk/handle/10283/3203>. The full GWAS summary statistics for the 23andMe discovery dataset will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Datasets will be made available at no cost for academic use. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information and to apply to access the data. GWAS summary statistics for brain structural measures are available at the Oxford Brain Imaging Genetics Server (<https://open.win.ox.ac.uk/ukbiobank/big40/>). According to the terms of consent, access to any form of individual-level data requires to approval and consent from each individual cohort.

REFERENCES

1. The Lancet Global Health. Mental health matters. *Lancet Glob Health*. 2020;8:e1352.
2. Pettersson E, Lichtenstein P, Larsson H, Song J, Agrawal A, Børglum AD, et al. Genetic influences on eight psychiatric disorders based on family data of 4 408 646 full and half-siblings, and genetic data of 333 748 cases and controls. *Psychol Med*. 2019;49:1166–73.
3. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22:343–52.
4. Shen X, Howard DM, Adams MJ, Hill WD, Clarke TK, Major Depressive Disorder Working Group of the Psychiatric Genomics C. et al. A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. *Nat Commun*. 2020;11:2301.
5. Schmaal L, Hibar DP, Samann PG, Hall GB, Baune BT, Jahanshad N, et al. Cortical abnormalities in adults and adolescents with Major Depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol Psychiatry*. 2017;22:900–9.
6. Shen X, Reus LM, Cox SR, Adams MJ, Lieuwald DC, Bastin ME, et al. Subcortical volume and white matter integrity abnormalities in Major Depressive Disorder: findings from UK Biobank imaging data. *Sci Rep*. 2017;7:5547.
7. Schmaal L, Veltman DJ, van Erp TG, Samann PG, Frodl T, Jahanshad N, et al. Subcortical brain alterations in Major Depressive Disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry*. 2016;21:806–12.
8. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of Major Depression. *Nat Genet*. 2018;50:668–81.
9. Thompson PM, Jahanshad N, Schmaal L, Turner JA, Winkler AM, Thomopoulos SI, et al. The enhancing neuroimaging genetics through meta-analysis consortium: 10 years of global collaborations in human brain mapping. *Hum Brain Mapp*. 2022;43:15–22.
10. Reus LM, Shen X, Gibson J, Wigmore E, Ligthart L, Adams MJ, et al. Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in UK Biobank. *Sci Rep*. 2017;7:42140.
11. Cattarinussi G, Delvecchio G, Sambataro F, Brambilla P. The effect of polygenic risk scores for major depressive disorder, bipolar disorder and schizophrenia on morphological brain measures: a systematic review of the evidence. *J Affect Disord*. 2022;310:213–22.
12. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*. 2022;603:654–60.
13. Tervo-Clemmens B, Marek S, Chauvin RJ, Van AN, Kay BP, Laumann TO, et al. Reply to: Multivariate BWAS can be replicable with moderate sample sizes. *Nature*. 2023;615:E8–E12.
14. Thompson PM, Jahanshad N, Ching CRK, Salminen LE, Thomopoulos SI, Bright J, et al. ENIGMA and global neuroscience: a decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry*. 2020;10:100.
15. Schmaal L, Pozzi E, Ho TC, van Velzen LS, Veer IM, Opel N, et al. ENIGMA MDD: seven years of global neuroimaging studies of Major Depression through worldwide data sharing. *Transl Psychiatry*. 2020;10:172.
16. Posner J, Biezonski D, Pieper S, Duarte CS. Genetic studies of mental illness: are children being left behind? *J Am Acad Child Adolesc Psychiatry*. 2021;60:672–4.
17. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 2016;19:1523–36.
18. Casey BJ, Cannonier T, Conley MJ, Cohen AO, Barch DM, Heitzeg MM, et al. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev Cogn Neurosci*. 2018;32:43–54.
19. Choi SW, O'Reilly PF. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience*. 2019;8:giz082.
20. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68–74.
21. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nat Commun*. 2019;10:5086.
22. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, et al. The genetic architecture of the human cerebral cortex. *Science*. 2020;367:eaay6690.
23. Harris MA, Cox SR, de Nooij L, Barbu MC, Adams MJ, Shen X, et al. Structural neuroimaging measures and lifetime depression across levels of phenotyping in UK biobank. *Transl Psychiatry*. 2022;12:157.
24. Furtjes AE, Cole JH, Couvy-Duchesne B, Ritchie SJ. A quantified comparison of cortical atlases on the basis of trait morphometricity. *Cortex*. 2023;158:110–26.
25. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002;33:341–55.
26. Nelder JA, Wedderburn RWM. Generalized linear models. *J R Stat Soc Ser A Stat Soc*. 2018;135:370–84.
27. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.
28. Therneau TM. *coxme: Mixed Effects Cox Models*. R package version 2.2-22. 2024; <https://CRAN.R-project.org/package=coxme>.
29. Saragosa-Harris NM, Chaku N, MacSweeney N, Guazzelli Williamson V, Scheuplein M, Feola B, et al. A practical guide for researchers and reviewers using the ABCD study and other large longitudinal datasets. *Dev Cogn Neurosci*. 2022;55:101115.
30. Shen X, Caramaschi D, Adams MJ, Walker RM, Min JL, Kwong A, et al. DNA methylome-wide association study of genetic risk for depression implicates antigen processing and immune responses. *Genome Med*. 2022;14:36.
31. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26:2867–73.
32. Alfaro-Almagro F, McCarthy P, Afyouni S, Andersson JLR, Bastiani M, Miller KL, et al. Confound modelling in UK Biobank brain imaging. *Neuroimage*. 2021;224:117002.
33. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67:1–48.
34. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601.
35. Brouwer RM, Klein M, Grasby KL, Schnack HG, Jahanshad N, Teeuw J, et al. Genetic variants associated with longitudinal changes in brain structure across the lifespan. *Nat Neurosci*. 2022;25:421–32.
36. Shen X, MacSweeney N, Chan SWY, Barbu MC, Adams MJ, Lawrie SM, et al. Brain structural associations with depression in a large early adolescent sample (the ABCD study(R)). *EClinicalMedicine*. 2021;42:101204.
37. Rolls ET, Cheng W, Feng J. The orbitofrontal cortex: reward, emotion and depression. *Brain Commun*. 2020;2:fcaa196.
38. Long Z, Duan X, Wang Y, Liu F, Zeng L, Zhao JP, et al. Disrupted structural connectivity network in treatment-naïve depression. *Prog Neuropsychopharmacol Biol Psychiatry*. 2015;56:18–26.

39. Fowler CH, Gaffrey MS. Reduced cortical surface area globally and in reward-related cortex is associated with elevated depressive symptoms in preschoolers. *J Affect Disord.* 2022;319:286–93.
40. Koenigs M, Huey ED, Calamia M, Raymond V, Tranel D, Grafman J. Distinct regions of prefrontal cortex mediate resistance and vulnerability to depression. *J Neurosci.* 2008;28:12341–8.
41. Rao VR, Sellers KK, Wallace DL, Lee MB, Bijanzadeh M, Sani OG, et al. Direct electrical stimulation of lateral orbitofrontal cortex acutely improves mood in individuals with symptoms of depression. *Curr Biol.* 2018;28:3893–902.e3894.
42. Ng TH, Alloy LB, Smith DV. Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit. *Transl Psychiatry.* 2019;9:293.
43. Tsuchida A, Doll BB, Fellows LK. Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. *J Neurosci.* 2010;30:16868–75.
44. Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? an fMRI study of social exclusion. *Science.* 2003;302:290–2.
45. Kross E, Berman MG, Mischel W, Smith EE, Wager TD. Social rejection shares somatosensory representations with physical pain. *Proc Natl Acad Sci USA.* 2011;108:6270–5.
46. Price JL. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci.* 2007;1121:54–71.
47. MacMaster FP, Kusumakar V. Hippocampal volume in early onset depression. *BMC Med.* 2004;2:2.
48. Sheline YI, Liston C, McEwen BS. Parsing the hippocampus in depression: chronic stress, hippocampal volume, and major depressive disorder. *Biol Psychiatry.* 2019;85:436–8.
49. Videbech P, Ravnkilde B. Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry.* 2004;161:1957–66.
50. Teicher MH, Samson JA. Annual research review: enduring neurobiological effects of childhood abuse and neglect. *J Child Psychol Psychiatry.* 2016;57:241–66.
51. Kraus C, Klobl M, Tik M, Auer B, Vanicek T, Geissberger N, et al. The pulvinar nucleus and antidepressant treatment: dynamic modeling of antidepressant response and remission with ultra-high field functional MRI. *Mol Psychiatry.* 2019;24:746–56.
52. Yang J, Hellerstein DJ, Chen Y, McGrath PJ, Stewart JW, Peterson BS, et al. Serotonin-norepinephrine reuptake inhibitor antidepressant effects on regional connectivity of the thalamus in persistent depressive disorder: evidence from two randomized, double-blind, placebo-controlled clinical trials. *Brain Commun.* 2022;4:fcc100.
53. Yamamura T, Okamoto Y, Okada G, Takaishi Y, Takamura M, Mantani A, et al. Association of thalamic hyperactivity with treatment-resistant depression and poor response in early treatment for major depression: a resting-state fMRI study using fractional amplitude of low-frequency fluctuations. *Transl Psychiatry.* 2016;6:e754.
54. Harris JA, Mihalas S, Hirokawa KE, Whitesell JD, Choi H, Bernard A, et al. Hierarchical organization of cortical and thalamic connectivity. *Nature.* 2019;575:195–202.
55. Amariuta T, Ishigaki K, Sugishita H, Ohta T, Koido M, Dey KK, et al. Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nat Genet.* 2020;52:1346–54.
56. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell.* 2019;179:589–603.

ACKNOWLEDGEMENTS

This work is supported by Wellcome Trust and UKRI grants to AMM (220857/Z/20/Z, 104036/Z/14/Z, 216767/Z/19/Z and MR/Z000548/1). GS also receives support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. XS is supported by UKRI award (ImmunoMIND) (Ref: MR/Z50354X/1). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

AUTHOR CONTRIBUTIONS

XS, HCW and AMM conceptualized and designed the study. XS and HCW contributed significantly to manuscript writing, and YJT, LKM and LS provided valuable feedback on the analyses and manuscript writing throughout the development of the project. NAG, SMK, LTS, UD, AJ, TK, IN, KS, BS, HV, DJS, SEM, KB, HJG, AK, KLM, GdZ, DJV, EP, SIT, NJ, PMT and LS were involved in data and funding acquisition and overseeing data preparation and analysis within each individual cohort. XS developed the analytical pipeline, and YJT and LKM tested and provided valuable feedback to optimise the code. EP, SIT, NJ, PMT and LS oversaw and organised data sharing and

analysis within the ENIGMA consortium. YJT, LKM, AW, NA, TFMA, KB, AJF, DG, TH, MH, NH, HJ, SM, YM, PGS, FS, AS, LT, GT, MJA, FTO, PU, UV, KW and MHR contributed substantially to data management, data acquisition and analysis for individual cohorts. YJ, CT and the 23andMe Research Team oversaw the analyses relevant to the use of GWAS summary statistics from 23andMe. All authors contributed to and critically reviewed the manuscript.

FUNDING:

ENIGMA MDD: ENIGMA MDD is supported by NIH (RO1 MH129832 Morey & Schmaal; RO1 MH129742 Thompson). LS is supported by an NHMRC Investigator Leadership Grant (2017962) and a University of Melbourne Dame Kate Campbell Fellowship. LH is funded by the Rubicon award (grant number 452020227) from the Dutch NWO. SEM is funded by the Australian NHMRC APP1158127 and APP1172917. YT is supported by Convergence | Healthy Start, a program of the Convergence Alliance – Delft University of Technology, Erasmus University Rotterdam, and Erasmus Medical Center. The ENIGMA-MDD working group gratefully acknowledges support from the NIH Big Data to Knowledge (BD2K) award (U54 EB020403 to Paul Thompson). For a complete list of ENIGMA-related grant support please see here: <http://enigma.ini.usc.edu/about-2/funding/>. **UK Biobank:** This research has been conducted using the UK Biobank Resource under Application Number 4844. **23andMe, Inc:** We would like to thank the research participants and employees of 23andMe for making this work possible. The following members of the 23andMe Research Team contributed to this study: Stella Aslibekyan, Adam Auton, Elizabeth Babalola, Robert K Bell, Jessica Bielenberg, Jonathan Bowes, Katarzyna Bryc, Ninad S Chaudhary, Daniella Coker, Sayantan Das, Emily DelloRusso, Sarah L Elson, Nicholas Eriksson, Teresa Filstein, Pierre Fontanillas, Will Freyman, Zach Fuller, Chris German, Julie M Granka, Karl Heilbron, Alejandro Hernandez, Barry Hicks, David A Hinds, Ethan M Jewett, Yunxuan Jiang, Katelyn Kukar, Alan Kwong, Yanyu Liang, Keng-Han Lin, Bianca A Llamas, Matthew H McIntyre, Steven J Micheletti, Meghan E Moreno, Priyanka Nandakumar, Dominique T Nguyen, Jared O'Connell, Aaron A Petrakovich, G David Poznik, Alexandra Reynoso, Shubham Saini, Morgan Schumacher, Leah Selcer, Anjali J Shastri, Janie F Shelton, Jingchunzi Shi, Suyash Shringarpure, Qiaojuan Jane Su, Susana A Tat, Vinh Tran, Joyce Y Tung, Xin Wang, Wei Wang, Catherine H Weldon, Peter Wilton, Corinna D Wong YJ and CT are employed by and hold stock or stock options in 23andMe, Inc. **FOR 2107:** This work was funded by DFG grants FOR2107 KI588/14-1, and KI588/14-2, and KI588/20-1, KI588/22-1 to Tilo Kircher, Marburg, Germany, grant FOR2107 DA1151/5-1, DA1151/5-2, DA1151/9-1, DA1151/10-1, DA1151/11-1 to UD; SFB-TRR58, Projects C09 and Z02 to UD; SFB/TRR 393, project grant no 521379614 to UD, TK, IN, AJF, HJ, BS, TH, NA, FS, AJ) and the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Münster (grant Dan3/022/22 to UD). Biosamples and corresponding data were sampled, processed and stored in the Marburg Biobank CBBMR. Igor Nenadić was supported by DFG (Deutsche Forschungsgemeinschaft) Grants NE2254/1-2, NE2254/2-1, NE2254/3-1, NE2254/4-1. Ethics approval was obtained from the ethics committees of the Medical Schools of the Universities of Marburg and Münster (Marburg: 07/2014, Münster: 2014-422-b-S), respectively, in accordance with the Declaration of Helsinki. All subjects volunteered to participate in the study and provided written informed consent. **MPiP:** The MPiP Munich Morphometry Sample comprises patients and control subjects included in Munich Antidepressant Response Signature study and the Recurrent Unipolar Depression (RUD) Case-Control study. We wish to acknowledge Rosa Schirmer, Elke Schreiter, Reinhold Borschke and Ines Eidner for image acquisition and data preparation, Anna Olynyk for data quality control, and Benno Pütz, Nazanin Karbalai, Darina Czamara, and Bertram Müller-Myhsok for distributed computing support. We thank Dorothee P Auer and Florian Holsboer for initiation of the RUD study. The MARS study was supported by a grant of the Exzellenz-Stiftung of the Max Planck Society. This work has also been funded by the BMBF in the framework of the National Genome Research Network (NGFN) (FKZ 01GS0481). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. **SHIP:** SHIP is part of the Community Medicine Research Network of the University Medicine Greifswald, which is supported by the German Federal State of Mecklenburg-West Pomerania. Genome-wide SNP typing in SHIP and MRI scans in SHIP and SHIP-TREND have been supported by a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. **BiDirect:** The BiDirect Study was supported by grants of the German Ministry of Research and Education (BMBF) to the University of Muenster (01ER0816 and 01ER1506). **QTIM:** QTIM was funded by the Australian National Health and Medical Research Council (Project Grants No. 496682 and 1009064) and US National Institute of Child Health and Human Development (RO1HD050735).

COMPETING INTERESTS

Yunxuan Jiang, Chao Tian are employed by and hold stock or stock options in 23andMe, Inc. Dr. Stein has received consultancy honoraria from Discovery Vitality,

Johnson & Johnson, Kanna, L'Oreal, Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. Hans J. Grabe has received travel grants and speakers' honoraria from Neuraxpharm, Servier, Indorsia and Janssen Cilag. PMT received partial research support from Biogen, Inc., for research unrelated to this manuscript. No other authors have conflict of interest to declare.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the NHS Tayside Research Ethics committee (05/s1401/89). Written consents were obtained from all participants from each individual ENIGMA cohort and publicly available datasets. All analyses and interpretation were performed under the guidance of the World Medical Association's Declaration of Helsinki (<https://www.wma.net/policies-post/wma-declaration-of-helsinki/>). The use, sharing and anonymisation for genetic and neuroimaging data were conducted under the EU General Data Protection Regulation (<https://gdpr-info.eu/>). Analyses and data management followed the guidelines of the UK Medical Research Council's 'Research, GDPR & Confidentiality' training. Access to UK Biobank individual-level was obtained under the application #4844. The ABCD study was approved by the National Institute of Mental Health Data Archive, United States (NIMH), and data was accessed through the NDA data base (Federal-Wide Assurance: FWA00018101).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41380-025-03136-4>.

Correspondence and requests for materials should be addressed to Xueyi Shen.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2025

¹Division of Psychiatry, University of Edinburgh, Edinburgh, UK. ²Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands. ³Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia. ⁴Orygen, Parkville, VIC, Australia. ⁵Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. ⁶Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany. ⁷German Center for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, Germany. ⁸Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. ⁹Department of Neurology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany. ¹⁰Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany. ¹¹Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany. ¹²Centre for Human Genetics, Philipps-University Marburg, Marburg, Germany. ¹³Institute for Translational Psychiatry, University of Münster, Münster, Germany. ¹⁴Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany. ¹⁵Institute of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany. ¹⁶Institute for Translational Neuroscience, University of Münster, Münster, Germany. ¹⁷Amsterdam Public Health, Mental Health program, Amsterdam, The Netherlands. ¹⁸Amsterdam Neuroscience, Complex Trait Genetics, Amsterdam, The Netherlands. ¹⁹Max Planck Institute of Psychiatry, Munich, Germany. ²⁰Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany. ²¹23andMe, San Francisco, CA, USA. ²²Department of Psychiatry and Neuroscience Institute, University of Cape Town, Cape Town, South Africa. ²³Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia. ²⁴West Region, Institute of Mental Health, Singapore, Singapore. ²⁵Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. ²⁶Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore. ²⁷Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. ²⁸SAMRC Research Unit on Risk and Resilience in Mental Disorders, University of Cape Town, Cape Town, South Africa. ²⁹Department of Psychiatry and Psychotherapy, University Hospital Bonn, Bonn, Germany. ³⁰School of Clinical Sciences, Queensland University of Technology, Brisbane, QLD, Australia. ³¹School of Psychology and Counselling, Queensland University of Technology, Brisbane, QLD, Australia. ³²Department of Psychiatry, Amsterdam University Medical Center, Amsterdam, The Netherlands. ³³Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ³⁴These authors jointly supervised this work: Andrew M. McIntosh, Heather C. Whalley. *Lists of authors and their affiliations appear at the end of the paper. [✉]email: Xueyi.shen@ed.ac.uk

THE 23ANDME RESEARCH TEAM

Yunxuan Jiang²¹ and Chao Tian²¹

FOR THE ENIGMA MDD WORKING GROUP

Xueyi Shen ^{1,✉}, Yara J. Toenders ², Laura K. M. Han ^{3,4,5}, Antoine Weihs ^{6,7}, Nina Alexander⁸, Till F. M. Andlauer ⁹, Katharina Brosch ⁸, Andreas J. Forstner ^{10,11,12}, Dominik Grotegerd¹³, Tim Hahn ¹³, Marco Hermesdorf¹⁴, Norbert Hosten¹⁵, Hamidreza Jamalabadi ⁸, Susanne Meinert ^{13,16}, Yuri Milaneschi ^{5,17,18}, Philipp G. Sämann ¹⁹, Frederike Stein⁸, Lea Teutenberg⁸, Florian Thomas-Odenthal ⁸, Paula Usemann⁸, Uwe Völker²⁰, Katharina Wittfeld ⁶, Marisol Herrera-Rivero ¹⁴, Nynke A. Groenewold ²², Sheri-Michelle Koopowitz²², Lachlan T. Strike²³, Udo Dannlowski ¹³, Andreas Jansen⁸, Tilo Kircher ⁸, Igor Nenadić ⁸, Kang Sim^{24,25,26}, Benjamin Straube ⁸, Henry Völzke²⁷, Dan J. Stein ^{22,28}, Sarah E. Medland ²³, Klaus Berger¹⁴, Hans J. Grabe⁶, Axel Krug²⁹, Katie L. McMahon ³⁰ and Greig de Zubicar³¹, Elena Pozzi³⁴, Dick J. Veltman³², Sophia I. Thomopoulos³³, Neda Jahanshad³³, Paul M. Thompson³³, Lianne Schmaal ^{3,4}, Andrew M. McIntosh ^{1,34} and Heather C. Whalley ^{1,34}