Supplementary Online Content

Winter NR, Leenings R, Ernsting J, et al. Quantifying deviations of brain structure and function in major depressive disorder across neuroimaging modalities. *JAMA Psychiatry*. Published online July 27, 2022. doi:10.1001/jamapsychiatry.2022.1780

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods 1. PubMed Search Terms

To get a rough estimate of the number of studies specifically investigating case-control differences between healthy and depressive subjects in neuroimaging modalities, we conducted a PubMed (https://pubmed.ncbi.nlm.nih.gov/advanced/) search on the 28th of September 2021 using the following search term which includes depression, a healthy control group and neuroimaging as key search components, resulting in a total of 1,585 publications:

(((major depressive disorder[Title]) OR (depression[Title]) OR (major depression[Title])) AND ((neuroimaging[Title/Abstract]) OR (magnetic resonance imaging[Title/Abstract]) OR (MRI[Title/Abstract])) AND ((healthy[Title/Abstract]) OR (control group[Title/Abstract]) OR (case control[Title/Abstract])))

eMethods 2. Exclusion Criteria

Subjects with any history of neurological (e.g., concussion, stroke, tumor, neuro-inflammatory diseases) and medical (e.g., cancer, chronic inflammatory or autoimmune diseases, heart diseases, diabetes mellitus, infections) conditions as well as non-Caucasian subjects were excluded from the analyses. Ethnicity was reported by the participants. Non-Caucasian subjects were excluded because the FOR2107 MACS cohort was originally recruited for the purpose of genetic and neuroimaging analyses, thus aiming to create a genetically homogeneous sample. The exclusion criteria were kept the same for depressive and healthy subjects. For HC, further exclusion criteria were current or history of any psychiatric illness.

eMethods 3. Diagnosis and Remission Status

A structured clinical interview for DSM-IV (SCID-I) was conducted with each participant in order to assess current and lifetime psychopathological diagnoses.[1] MDD diagnosis was based on the criteria defined in DSM-IV which requires the presence of five of nine cardinal symptoms for two weeks or longer within the last 4 weeks, are present for most of the day nearly every day and need to cause significant distress or impairment. To fulfill the diagnostic criteria, a depressed mood or markedly diminished interest or pleasure have to be present (at least one or both). Other symptoms are clinically significant weight gain/loss or appetite disturbance, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue or loss of energy, feelings of worthlessness or excessive guilt, diminished ability to concentrate or think clearly, and recurrent thoughts of death or suicide. Only participants with a primary MDD diagnosis were included in the MDD sample. MDD patients were not excluded if they fulfilled the criteria of an additional comorbid psychopathological diagnosis. Partial remission was defined either by (1) a presence of some major depressive symptoms but full criteria are no longer met or (2) no major depressive symptoms but the period of remission has been less than two months. Complete remission was defined as the disappearance of the diagnostic criteria of depression for at least two consecutive months.

eMethods 4. Calculation of Medication Load Index

A medication load index was calculated expressing the level of current psychiatric medication in terms of dosage of the combined current medication. Assessment of acute medication was done based on interviews conducted by trained personnel. Following the interview, all recorded medication was then classified by the corresponding

active ingredients and medication categories (e.g. selective serotonin reuptake inhibitors, benzodiazepines; somatic medication was grouped in categories as contraceptives, thyroid or diabetes medication). Subsequently, all active ingredients of psychiatric medication were coded '0', '1' or '2' based on established dosage-dependent cutoffs.[2] These scores were then added together across all psychiatric medication to results into the final medication index. This procedure has been used in numerous previous publications.[3,4]

eMethods 5. Analysis of the Depressive Subgroups

For the secondary analyses including only acutely depressed patients, we removed fully remitted patients from the MDD sample. For the secondary analyses including only chronically depressed patients, we only used MDD subjects that had a history of at least two inpatient stays due to their depressive disorder. To explore the effect of current antidepressant or other psychotropic medication on group differences between healthy and depressive subjects, we analyzed a subgroup of depressive patients that were medicated during the time of the scan. For this, we excluded patients with a medication load index of 0. From the full depressive sample (861), 341 patients with no current medication were excluded for this analysis. Within the remaining sample of 520 patients, medication consisted of selective serotine-norepinephrine reuptake inhibitor (39.0%), selective serotonin reuptake inhibitor (41.7%), norepinephrine-dopamine reuptake inhibitor (5.2%), noradrenergic and specific serotonergic antidepressant (13.5%), noradrenaline reuptake inhibitor (0.2%), tricyclic antidepressants (7.5%), monoamine oxidase inhibitors (0.8%), agomelatine (6.7%), lithium (3.7%), benzodiazepine (1.2%), z-drugs (1.2%), and stimulants (0.8%).

eMethods 6. Matching Procedure for an Age- and Sex-Matched Healthy Sample

In order to assess whether matching a healthy to the depressive sample based on sex and age, we used the *R* package *MatchIt*.[7] *MatchIt* implements a wide range of matching methods for improving parametric statistical models. We used *R* version 4.1.0 and *MatchIt* version 4.3.3. We used default parameters for the matching procedure, i.e., a nearest neighbor matching on the propensity score estimated using logistic regression. We aimed for a ratio of 1:1 healthy and depressive subjects. The matching was based on sex and age. The results of the matched samples analyses are reported in the supplementary results section.

eMethods 7. Statistical Covariates and Correction

As the MRI body-coil was changed mid recruitment in Marburg, scanning site was dummy coded to represent three categories (pre and post body-coil change Marburg and Münster) and used in the Hariri task analysis. For PRS analyses, the first three Multi-Dimensional Scaling (MDS) components of the genetic relationship matrix were added as covariates to correct for population stratification (for details, see [5]).

For all but voxel-wise analyses, we controlled for multiple comparisons by calculating the false discovery rate with a false positive rate of 0.05 using the Benjamini-Yekutieli procedure.[6] For voxel-based data, significance thresholds for multiple testing were obtained at the voxel-level using non-parametric *t* tests as implemented in the SPM Threshold-Free Cluster Enhancement (TFCE) toolbox. An FWE-corrected threshold of 0.05 was used to calculate corrected p-values. We decided against a cluster correction for voxel-based analyses as implemented in the TFCE toolbox since the peak voxel of a significant cluster must not necessarily show the strongest effect among all voxels in the brain. This is, of course, usually a desirable characteristic of the TFCE correction. However, as we intended to find the upper bound, i.e., the largest difference between healthy and depressive subjects, we used a correction method on a voxel-level.

The predictive utility estimates (accuracy, sensitivity, specificity, and AUROC) presented in the manuscript are based on in-sample statistics. We did not compute cross-validated metrics to 1) mirror current practice in the field and 2) estimate the upper bounds for the predictive utility. This way, our results represent the maximum predictive utility that can be expected in similar analyses. Cross-validated out-of-sample accuracies as a measure of generalizability of predictive markers will fall below the upper bound that we present in this manuscript.

eMethods 8. Assessment of Childhood Maltreatment and Social Support

The well-established childhood trauma questionnaire (CTQ) was used to assess childhood maltreatment in patients and controls.[8] For all analyses including childhood maltreatment, we used a sum score that can be computed across the five maltreatment subscales. For 12 subjects no CTQ sum score was available, resulting in a sample of 1797 subjects that could be used for the CTQ analysis.

Perceived social support was measured using the Social Support Questionnaire (F-SozU), an established German self-report instrument.[9] The questionnaire consists of three subscales measuring the subject's perceived emotional support, instrumental support, and social integration. A sum score aggregating all three scales was used in all analyses. For 10 subjects no social support sum score was available, resulting in a sample of 1799 subjects that could be used for the social support analysis.

eMethods 9. Polygenic Risk Scores

Genotyping was conducted using the PsychArray BeadChip, followed by quality control and imputation, as described previously.[5,10,11] In short, QC and population substructure analyses were performed in PLINK v1.9.[12] The data were imputed to the 1000 Genomes phase 3 reference panel using SHAPEIT and IMPUTE2.[13,14] Imputed genetic data were available for n=1689 individuals. Genetically related participants were identified in PLINK using the command –genome and one individual of each related pair (PI-HAT ≥12.5) was excluded for the specific analysis (e.g., MDD subgroups). For the main HC versus MDD analysis, 68 related participants were excluded from the sample, resulting in a final sample of n=1621. Excluding related participants from the whole genetic sample (n=1689) was done separately for every analysis, e.g., when comparing only acutely depressed patients with healthy subjects, to guarantee that only a minimal number of related samples were excluded for every MDD, sex, and site subgroup analysis. Sample sizes for all other analyses are listed in *eTable 2-4*.

From the genetic data, a polygenic risk score was calculated using the PRS-CS[15] method using summary statistics from a recent MDD GWAS.[16] PRS-CS uses Bayesian regression to infer PRS weights and models the local linkage disequilibrium pattern of all variants. To this end, PRS-CS uses a global scaling parameter phi. We estimated this parameter automatically using the PRS-CS-auto method at $\varphi=1\cdot30\times10^{-4}$. This results in only a single PRS that we used for our analysis, in contrast to polygenic risk scores based on p-thresholds that will return one score per p-threshold. The FOR2107 MACS data used in this study was independent from the MDD GWAS.

To control for population substructure, three MDS components were added as covariates to all linear models containing genetic data. The MDS components were calculated separately for every analysis, e.g., when comparing only acutely depressed patients with healthy subjects.

eMethods 10. Magnetic Resonance Imaging
Magnetic resonance imaging (MRI) data for all brain-based modalities were acquired using two 3T whole body MRI scanner (Marburg: MAGNETOM Trio Tim, software version Syngo MR B17, Siemens, Erlangen, Germany, 12-channel head matrix Rx-coil; Münster: MAGNETOM Prisma, software version Syngo MR D13D, Siemens, Erlangen, Germany, 20-channel head matrix Rx-coil) using a GRAPPA acceleration factor of 2. Pulse sequence parameters were standardized across both sites to the extent permitted by each platform (see below for differences between sites). Due to a change of the scanner body-coil in Marburg during recruitment, three different scanner conditions were used as covariate in all MRI-based analyses (Marburg pre body-coil change, Marburg post body-coil change and Münster. Detailed information on pulse sequence parameters as well as quality assurance protocols have been described previously in [17].

eMethods 11. T1-Weighted MRI
Structural MRI data was acquired using the following parameters: TE = 2.26ms (Marburg), TE = 2.28ms (Münster), TR = 1,900ms (Marburg), TR = 2,130ms (Münster), FoV = 256 mm, matrix = 256 × 256, slice thickness = 1 mm, distance factor = 50%, phase encoding direction anterior >> posterior, flip angle = 8°, bandwidth 200 Hz/Px, ascending acquisition, axial acquisition, 192 slices. There was no T_1 scan available for 18 subjects (remaining sample n=1791). Image quality was assessed by visual inspection of a trained expert as well as by checking for image homogeneity using the CAT12 toolbox. 47 subjects were excluded due to low image quality and artifact (remaining sample n = 1744). Three additional subjects were excluded from the Freesurfer analysis (n=1741) due to poor segmentation quality (see below).

Voxel-based morphometry (VBM) was preprocessed using the CAT12 MATLAB toolbox (version r1450) with default parameters.[18] In short, images were bias-corrected, tissue classified, and normalized to MNI-space using linear and non-linear transformations. Normalization was done using a pre-computed high-dimensional DARTEL template. The modulated gray matter images were smoothed with a Gaussian kernel of 8 mm FWHM. Absolute threshold masking with a threshold value of 0.1 was used for all analyses as recommended in the CAT12 manual (http://www.neuro.uni-jena.de/cat/).

Automated segmentation was conducted using the cortical and subcortical parcellation stream of Freesurfer (Version 5.3) based on the Desikan-Killiany atlas.[19] In total, measures for 68 cortical regions (34 on each hemisphere), 14 subcortical regions (7 on each hemisphere), 4 ventricles (2 on each hemisphere), as well as total intracranial volume (ICV) were extracted for each participant. Additionally, global measures of cortical and subcortical surface, thickness and volume were calculated per and across hemisphere (for a full list of used Freesurfer parameters, see *Supplementary Methods 10*). This resulted in a total number of 166 parameters that were used in the statistical analyses. Default parameters were used for the segmentation (https://surfer.nmr.mgh.harvard.edu/) and segmentation quality was reviewed visually as well as based on statistical outlier analysis following standardized protocols by the ENIGMA consortium.[20] After excluding subjects with poor segmentation quality, a final sample of 1741 subjects were available for the Freesurfer analysis. Missing values for single regions were median imputed.

eMethods 12. Freesurfer Measures
L_bankssts_thickavg

L_caudalanteriorcingulate_thickavg

L_caudalmiddlefrontal_thickavg

L_cuneus_thickavg

L_entorhinal_thickavg

L_fusiform_thickavg

L_inferiorparietal_thickavg

L_inferiortemporal_thickavg

L_isthmuscingulate_thickavg

L_lateraloccipital_thickavg

L_lateralorbitofrontal_thickavg

L_lingual_thickavg

L_medialorbitofrontal_thickavg

L_middletemporal_thickavg

L_parahippocampal_thickavg

L_paracentral_thickavg

L_parsopercularis_thickavg

L_parsorbitalis_thickavg

L_parstriangularis_thickavg

L_pericalcarine_thickavg

L_postcentral_thickavg

L_posteriorcingulate_thickavg

L_precentral_thickavg

L_precuneus_thickavg

L_rostralanteriorcingulate_thickavg

L_rostralmiddlefrontal_thickavg

L_superiorfrontal_thickavg

L_superiorparietal_thickavg

L_superiortemporal_thickavg

L_supramarginal_thickavg

L_frontalpole_thickavg

L_temporalpole_thickavg

L_transversetemporal_thickavg

L_insula_thickavg

R_bankssts_thickavg

R_caudalanteriorcingulate_thickavg

R_caudalmiddlefrontal_thickavg

R_cuneus_thickavg

R_entorhinal_thickavg

R_fusiform_thickavg

R_inferiorparietal_thickavg

R_inferiortemporal_thickavg

R_isthmuscingulate_thickavg

R_lateraloccipital_thickavg

R_lateralorbitofrontal_thickavg

R_lingual_thickavg

R_medialorbitofrontal_thickavg

R_middletemporal_thickavg

R_parahippocampal_thickavg

R_paracentral_thickavg

R_parsopercularis_thickavg

R_parsorbitalis_thickavg

R_parstriangularis_thickavg

R_pericalcarine_thickavg

R_postcentral_thickavg

R_posteriorcingulate_thickavg

R_precentral_thickavg

R_precuneus_thickavg

R_rostralanteriorcingulate_thickavg

R_rostralmiddlefrontal_thickavg

R_superiorfrontal_thickavg

R_superiorparietal_thickavg

R_superiortemporal_thickavg

R_supramarginal_thickavg

R_frontalpole_thickavg

R_temporalpole_thickavg

R_transversetemporal_thickavg

R_insula_thickavg

LThickness

RThickness

LSurfArea

RSurfArea

L_bankssts_surfavg

L_caudalanteriorcingulate_surfavg

L_caudalmiddlefrontal_surfavg

L_cuneus_surfavg

L_entorhinal_surfavg

L_fusiform_surfavg

L_inferiorparietal_surfavg

L_inferiortemporal_surfavg

L_isthmuscingulate_surfavg

L_lateraloccipital_surfavg

L_lateralorbitofrontal_surfavg

L_lingual_surfavg

L_medialorbitofrontal_surfavg

L_middletemporal_surfavg

L_parahippocampal_surfavg

L_paracentral_surfavg

L_parsopercularis_surfavg

L_parsorbitalis_surfavg

L_parstriangularis_surfavg

L_pericalcarine_surfavg

L_postcentral_surfavg

L_posteriorcingulate_surfavg

L_precentral_surfavg

L_precuneus_surfavg

L_rostralanteriorcingulate_surfavg

L_rostralmiddlefrontal_surfavg

L_superiorfrontal_surfavg

L_superiorparietal_surfavg

L_superiortemporal_surfavg

L_supramarginal_surfavg

L_frontalpole_surfavg

L_temporalpole_surfavg

L_transversetemporal_surfavg

L_insula_surfavg

R_bankssts_surfavg

R_caudalanteriorcingulate_surfavg

R_caudalmiddlefrontal_surfavg

R_cuneus_surfavg

R_entorhinal_surfavg

R_fusiform_surfavg

R_inferiorparietal_surfavg

R_inferiortemporal_surfavg

R_isthmuscingulate_surfavg

R_lateraloccipital_surfavg

R_lateralorbitofrontal_surfavg

R_lingual_surfavg

R_medialorbitofrontal_surfavg

R_middletemporal_surfavg

R_parahippocampal_surfavg

R_paracentral_surfavg

R_parsopercularis_surfavg

R_parsorbitalis_surfavg

R_parstriangularis_surfavg

R_pericalcarine_surfavg

R_postcentral_surfavg

R_posteriorcingulate_surfavg

R_precentral_surfavg

R_precuneus_surfavg

R_rostralanteriorcingulate_surfavg

R_rostralmiddlefrontal_surfavg

R_superiorfrontal_surfavg

R_superiorparietal_surfavg

R_superiortemporal_surfavg

R_supramarginal_surfavg

R_frontalpole_surfavg

R_temporalpole_surfavg

R_transversetemporal_surfavg

R_insula_surfavg

ICv

LLatVent

RLatVent

Lthal

Rthal

Lcaud

Rcaud

Lput

Rput

Lpal

Rpal

Lhippo

Rhippo

Lamyg

Ramyg

Laccumb

Raccumb

BrainSegVol

BrainSegVolNotVent

lhCortexVol

rhCortexVol

CortexVol

lh.WhiteMatterVol

rh.WhiteMatterVol

SubCortGrayVol

TotalGrayVol

eMethods 13. Functional MRI Image Acquisition

For the two functional MRI paradigms (face matching, resting-state), a T_2*-weighted echo-planar imaging (EPI) sequence sensitive to blood oxygen level-dependent (BOLD) contrast was used with the following parameters: TE = 30 ms (Marburg), TE = 29ms (Münster), TR = 2,000 ms, FoV = 210 mm, matrix = 64 × 64, slice thickness = 3.8 mm, distance factor = 10%, phase encoding direction anterior >> posterior, flip angle = 90°, no parallel imaging, bandwidth 2,232 Hz/Px, ascending acquisition, axial acquisition, 33 slices, slice alignment parallel to AC-PC line tilted 20° in the dorsal direction.

For resting-state fMRI, two-hundred-thirty-seven interleaved and ascending measurements (8 minutes) tilted with -20° against the anterior and posterior commission alignment (AC-PC alignment) were acquired (Base resolution: 64, Bandwith: 2232 Hz/Px, Echo spacing: 0.51ms EPI factor: 64, BOLD threshold: 4, TR: 2s). Participants were asked to keep their eyes closed until the end of the resting state session.

eMethods 14. Resting-State fMRI

Participants were asked to keep their eyes closed until the end of the resting state session. Resting-state preprocessing was done using the CONN (v18b) MATLAB toolbox and the default volume-based MNI preprocessing pipeline.[21] First, a series of preprocessing steps was applied to the functional and structural images that included a functional realignment and unwarp, a slice-timing correction, an ART-based outlier identification, a direct segmentation and normalization of the functional and structural images, as well as a final functional smoothing with an 8mm FWHM kernel. Second, CONN's denoising step with default parameters was applied to regress out potential noise artefacts in the functional data based on an anatomical component-based noise correction procedure (aCompCor).[22] Controlling for noise artefacts is done by regressing out noise components from cerebral white matter and cerebrospinal areas (5 PCA components), estimated subject-motion parameters (12 parameters including 6 motion parameters and their associated first-order derivatives), as well as identified outlier scans or scrubbing. Finally, temporal band-pass filtering was applied to remove low frequencies under 0.008 Hz and high frequencies above 0.09 Hz.

Resting-state and T_1 data was available for 1374 subjects. 12 subjects were excluded because of an acute medication with tranquilizers such as benzodiazepine or Z-drugs which are known to influence functional MRI data. 17 subjects were excluded after visual quality checks if functional or structural segmentation and normalization quality was poor or distribution of correlation values after denoising weren't following a normal distribution (using quality assurance plots in the CONN toolbox). 9 subjects were excluded due to strong motion artifacts that resulted in less than 5 minutes of valid resting-state image time points after scrubbing. Data of a final sample of 1336 subjects was available for resting-state analyses.

Connectivity matrices for each subject were created by computing the bivariate Pearson's correlation coefficient between the time-series of every region of the 17 networks Schaefer atlas with 100 parcels.[23] Amplitude of Low-Frequency Fluctuations (ALFF) maps were calculated from resting-state time-series. ALFF represents a measure of BOLD signal power within the frequency band of interest and is defined as the root mean square of BOLD signal at each voxel after filtering.[24] Fractional ALFF (fALFF) as a relative measure of BOLD signal power was calculated as the ratio of root mean square of BOLD signal at each individual voxel after vs. before low- or

band- pass filtering.[25] As a measure of regional homogeneity (ReHo), we have used the Local Correlation (LCOR) method implemented in the CONN toolbox. LCOR can be considered a generalization of the original ReHo definition by Zang et al. (2004) that is more robust across different data resolutions and neighboring sizes. LCOR is defined as the average of correlation coefficients between each individual voxel and a region of neighboring voxels.[21] Please see Deshpande et al. (2009) for a detailed description of the Integrated Local Correlation and a comparison to ReHo.[26]

eMethods 15. Task-Based fMRI

Functional MRI data from a well-established emotional face matching paradigm was used.[27] The experimental setup and preprocessing have been described previously.[28] In short, subjects viewed images of fearful or angry faces in the experimental and geometric shapes in the control condition. In each trial, a target image was presented at the top while two further images were presented at the bottom left and right, whereby one of these images was identical to the target image. The subjects were instructed to indicate whether the left or right image was identical to the top image by pressing a corresponding button.

Image acquisition details corresponded to the previously described resting-state scanning parameters. To avoid motion artifacts, subjects were excluded from the final sample if their overall movement exceeded 2mm. Additionally, a visual quality check has been conducted to exclude subjects with visually detectable artifacts. As described above, subjects under acute medication with tranquilizers have been excluded from the analyses. At the individual subject level, fMRI responses for both conditions (faces, shapes) were modeled in a block design using the canonical hemodynamic response function implemented in SPM8 convolved with a vector of onset times for the different stimulus blocks. High-pass filtering was applied with a cut-off frequency of 1/128 Hz to attenuate low-frequency components. Contrast images were created by contrasting beta images of the 'faces' against the 'shapes' condition. Face matching fMRI paradigm was available for 1368 participants. 4 subjects under acute medication with tranquilizers were excluded (remaining sample = 1364). 109 subjects with more than 2mm of movement were excluded (remaining sample = 1255). 14 subjects with low image quality or artifacts (after visual inspection of a trained expert) have been excluded. A final sample of 1241 subjects were available for the face matching task-based fMRI analyses.

eMethods 16. DTI Image Acquisition and Preprocessing

Image acquisition

Data were acquired using a GRAPPA acceleration factor of 2. Fifty-six axial slices with no gap were measured with an isotropic voxel size of 2.5 × 2.5 × 2.5 mm^3 (TE = 90 ms, TR = 7300 ms). Five non-DW images ($b = 0$ s/mm^2) and 2 × 30 DW images with a b-value of 1000 s/mm^2 were acquired.

Preprocessing

DTI image preprocessing was done as described in [29]. First, FSL's eddy was used to realign the Diffusion-weighted images (DWI) and correct those for eddy currents and susceptibility distortions.[30–33] Second, the CATO toolbox was used to reconstruct the anatomical connectome of the diffusion tensor imaging data (DTI) which models the measured signal of a single voxel by a tensor describing the preferred diffusion direction per voxel.[34] It makes use of the RESTORE algorithm which estimates the diffusion tensor while simultaneously identifying and removing outliers, thereby reducing the impact of physiological noise artifacts.[32,33] Third, the resulting

diffusion profiles were used to reconstruct white matter paths using deterministic tractography. To this end, eight seeds were started per voxel, and for each seed, a tractography streamline was constructed by following the main diffusion direction from voxel to voxel. Stop criteria included reaching a voxel with a fractional anisotropy < 0.1, making a sharp turn of >45°, reaching a gray matter voxel, or exiting the brain mask. Given the poorer DWI signal-to-noise ratio in subcortical regions and the dominant effect of subcortical regions on network properties, we decided to use a subdivision of FreeSurfer's Desikan Killiany Atlas containing only cortical regions, as we have done in previous work.[29,35–37]

For every subject, the network information was stored in a structural connectivity matrix, with rows and columns reflecting 114 cortical brain regions of the Lausanne parcellation, a subdivision of the FreeSurfer's Desikan-Killiany atlas.[35,36] Matrix entries represent the weights of the graph edges. Network edges were weighted according to fractional anisotropy (FA), mean diffusivity (MD), and number of streamlines (NOS). For all FA- and MD-based connectivity analyses, only edges with non-zero values for at least 95% of subjects were analyzed. This resulted in a varying number of subjects that were included in the statistical modeling of group differences in edge values. DTI data was available for 1567 subjects. 55 subjects were excluded after quality control (see below). A final sample of 1508 subjects were available for all DTI-based analyses.

Quality control

Four metrics were included in the detection of outliers: the average number of streamlines, the average fractional anisotropy, the average prevalence of each subject's connections (low value if the subject has "odd" connections), and the average prevalence of each subjects connected brain regions (high value, if the subject misses commonly found connections). Then, quartiles (Q1, Q2, Q3) and the interquartile range (IQR=Q3-Q1) was computed for every metric across the group. A datapoint was declared an outlier if its value was below Q1-1.5*IQR or above Q3+1.5*IQR on any of the four metrics.

eMethods 17. Graph Network Parameters

For the DTI connectivity matrix, a binary adjacency matrix was calculated on the basis of number of streamlines. All edges with less than three number of streamlines were set to 0, all other edges were set to 1. For the resting-state fMRI connectivity matrix, a binary adjacency matrix was calculated by setting the top 15 percent of connections (highest correlation coefficient) to 1, all other edges were set to 0. The adjacency matrices were then used to calculate a number of representative graph parameters using PHOTONAI Graph (https://github.com/wwu-mmll/photonai_graph), which itself calls function from the Python package network.[38] Used graph metrics were defined as follows (for an introduction of graph metrics for brain connectivity, see [39]): Global efficiency was defined as the average inverse shortest path length between all node pairs. Local efficiency was defined as the global efficiency on node neighborhoods. Clustering coefficient was computed as the average likelihood that the neighbors of a node are also mutually connected. Degree centrality was defined as the number of nodes connected to the node of interest. Betweenness centrality was defined as the fraction of shortest paths of the network that pass through the node of interest. Degree assortativity was defined as the similarity of connections in the graph with respect to the node degree. Clustering coefficient, degree centrality, and betweenness centrality were calculated per node and additionally averaged across nodes, resulting in a total of 348 network parameters.

eTable 1. Statistical Analyses of Differences Between Healthy and Acutely Depressed Individuals

HC versus Acutely Depressed MDD	n (HC, MDD)	*df1*	*df2*	*F*	p_{uncorr}	p_{corr}	$\eta^2_{partial}$	overlap, %	BACC, %	sens, %	spec, %	AUROC, %	covariates
Structural MRI													
VBM	1497 (926, 571)	1	1490	27.25	<0.001	0.056	0.018 [0.006 - 0.033]	89.15	55.75	54.75	56.74	58.32	a
Cortical and subcortical surface, thickness, volume	1492 (923, 569)	1	1486	14.62	<0.001	0.070	0.009 [0.002 - 0.023]	91.26	53.85	51.46	56.24	55.63	b
Task-based functional MRI													
Face matching task	1062 (656, 406)	1	1056	13.11	<0.001	0.628	0.012 [0.002 - 0.027]	90.97	54.83	53.51	56.16	56.44	b
Functional connectome													
Bivariate connectivity	1143 (702, 441)	1	1137	24.15	<0.001	0.046	0.021 [0.008 - 0.039]	87.18	55.25	54.27	56.24	58.18	b
Network parameters	1143 (702, 441)	1	1137	16.59	<0.001	0.095	0.014 [0.005 - 0.030]	89.97	54.51	50.28	58.73	56.72	b
Local Correlation	1143 (702, 441)	1	1137	19.66	<0.001	0.196	0.017 [0.006 - 0.035]	89.43	55.69	53.56	57.82	56.91	b
ALFF	1143 (702, 441)	1	1137	20.72	<0.001	0.101	0.018 [0.006 - 0.037]	86.21	54.82	61.11	48.53	56.79	b
fALFF	1143 (702, 441)	1	1137	20.76	<0.001	0.082	0.018 [0.006 - 0.036]	89.22	55.37	54.27	56.46	58.10	b
Structural connectome													
FA	1291 (809, 482)	1	1285	6.68	0.010	1.000	0.005 [0.000 - 0.016]	94.06	53.87	50.68	57.05	54.23	b
MD	1298 (814, 484)	1	1292	12.20	<0.001	0.679	0.009 [0.002 - 0.022]	89.52	54.10	47.05	61.16	55.44	b
Network parameters	1305 (819, 486)	1	1299	11.49	<0.001	1.000	0.009 [0.001 - 0.022]	91.48	54.05	67.16	40.95	55.58	b
Genetics													
Polygenic Risk Score	1401 (854, 547)	1	1393	53.11	<0.001	<0.001	0.037 [0.020 - 0.056]	84.25	59.12	57.73	60.51	61.11	c
Environment													
Social Support	1537 (945, 592)	1	1532	619.07	<0.001	<0.001	0.288 [0.247 - 0.329]	50.14	74.46	80.85	68.07	81.26	d
Childhood Maltreatment	1541 (947, 594)	1	1536	436.72	<0.001	<0.001	0.221 [0.186 - 0.255]	52.43	72.10	81.41	62.79	77.26	d

BACC = balanced accuracy, AUROC = area under the receiver operating curve, HC = Healthy Controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, FA = fractional anisotropy, MD = mean diffusivity. Covariates in the statistical models: a= age + sex + dummy scanner + total intracranial volume, b= age + sex + dummy scanner, c= age + sex + dummy site + MDS 1 + MDS 2 + MDS 3, d= age + sex + dummy site

eTable 2. Statistical Analyses of Differences Between Healthy and Chronically Depressed Individuals

HC versus Chronically Depressed MDD	n (HC, MDD)	*df1*	*df2*	*F*	p_{uncorr}	p_{corr}	$\eta^2_{partial}$	overlap, %	BACC, %	sens, %	spec, %	AUROC, %	covariates
Structural MRI													
VBM	1210 (926, 284)	1	1203	28.50	<0.001	0.016	0.023 [0.009 - 0.042]	86.22	55.22	55.51	54.93	58.56	a
Cortical and subcortical surface, thickness, volume	1207 (923, 284)	1	1201	21.19	<0.001	0.003	0.017 [0.005 - 0.032]	87.96	56.42	52.98	59.86	58.45	b
Task-based functional MRI													
Face matching task	843 (656, 187)	1	837	14.32	<0.001	0.657	0.017 [0.004 - 0.035]	87.71	55.21	55.34	55.08	58.46	b
Functional connectome													
Bivariate connectivity	910 (702, 208)	1	904	22.36	<0.001	0.118	0.024 [0.008 - 0.050]	85.68	57.43	58.12	56.73	59.83	b
Network parameter	910 (702, 208)	1	904	15.64	<0.001	0.158	0.017 [0.004 - 0.037]	87.85	56.51	54.84	58.17	58.74	b
Local Correlation	910 (702, 208)	1	904	24.69	<0.001	0.036	0.027 [0.010 - 0.052]	84.96	58.00	59.26	56.73	60.82	b
ALFF	910 (702, 208)	1	904	26.77	<0.001	0.027	0.029 [0.010 - 0.060]	79.07	56.33	64.10	48.56	58.71	b
fALFF	910 (702, 208)	1	904	23.14	<0.001	0.139	0.025 [0.009 - 0.050]	85.46	59.01	58.40	59.62	60.42	b
Structural connectome													
FA	1053 (819, 234)	1	1047	7.80	0.005	1.000	0.007 [0.001 - 0.021]	92.00	53.39	55.07	51.71	55.31	b
MD	1053 (819, 234)	1	1047	12.88	<0.001	0.481	0.012 [0.003 - 0.028]	88.29	55.25	50.67	59.83	56.72	b
Network parameter	1053 (819, 234)	1	1047	15.15	<0.001	0.236	0.014 [0.005 - 0.031]	87.85	55.10	52.50	57.69	57.88	b
Genetics													
Polygenic Risk Score	1132 (859, 273)	1	1124	35.16	<0.001	<0.001	0.030 [0.015 - 0.055]	83.94	59.21	57.97	60.44	61.51	c
Environment													
Social Support	1240 (945, 295)	1	1235	416.61	<0.001	<0.001	0.252 [0.200 - 0.306]	51.98	73.36	80.63	66.10	78.30	d
Childhood Maltreatment	1243 (947, 296)	1	1238	472.69	<0.001	<0.001	0.276 [0.229 - 0.325]	47.13	74.58	83.63	65.54	80.02	d

BACC = balanced accuracy, AUROC = area under the receiver operating curve, HC = Healthy Controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, FA = fractional anisotropy, MD = mean diffusivity. Covariates in the statistical models: a= age + sex + dummy scanner + total intracranial volume, b= age + sex + dummy scanner, c= age + sex + dummy site + MDS 1 + MDS 2 + MDS 3, d= age + sex + dummy site

eTable 3. Statistical Analyses of Differences Between Healthy and Medicated Depressive Individuals

HC versus Medicated MDD	n (HC, MDD)	*df1*	*df2*	*F*	p_{uncorr}	p_{corr}	$\eta^2_{partial}$	overlap, %	BACC, %	sens, %	spec, %	AUROC, %	covariates
Structural MRI													
VBM	1421 (926, 495)	1	1414	29.57	<0.001	0.018	0.020 [0.008 - 0.037]	88.19	54.97	55.18	54.75	57.92	a
Cortical and subcortical surface, thickness, volume	1417 (923, 494)	1	1411	21.11	<0.001	0.004	0.015 [0.005 - 0.029]	89.84	54.98	51.46	58.50	56.73	b
Task-based functional MRI													
Face matching task	995 (655, 340)	1	989	12.12	<0.001	0.816	0.012 [0.002 - 0.029]	91.03	54.16	54.20	54.12	55.68	b
Resting-state fMRI													
Bivariate connectivity	1080 (702, 378)	1	1074	27.13	<0.001	0.010	0.025 [0.010 - 0.048]	86.69	57.26	53.42	61.11	59.48	b
Network parameter	1080 (702, 378)	1	1074	16.08	<0.001	0.054	0.015 [0.004 - 0.033]	88.53	55.58	46.87	64.29	57.25	b
Local Correlation	1080 (702, 378)	1	1074	30.36	<0.001	0.021	0.027 [0.011 - 0.050]	85.55	56.86	58.69	55.03	59.01	b
ALFF	1080 (702, 378)	1	1074	21.82	<0.001	0.020	0.020 [0.006 - 0.040]	84.64	54.63	62.96	46.30	57.54	b
fALFF	1080 (702, 378)	1	1074	27.02	<0.001	0.020	0.025 [0.010 - 0.045]	87.21	56.03	55.98	56.08	58.73	b
Structural connectome													
FA	1238 (819, 419)	1	1232	6.60	0.010	1.000	0.005 [0.000 - 0.017]	93.93	52.10	54.09	50.12	53.96	b
MD	1191 (799, 392)	1	1185	10.32	0.001	1.000	0.009 [0.002 - 0.019]	88.03	52.17	46.43	57.91	54.09	b
Network parameter	1238 (819, 419)	1	1232	8.55	0.003	1.000	0.007 [0.001 - 0.017]	92.82	53.50	50.67	56.32	55.10	b
Genetics													
Polygenic Risk Score	1332 (855, 477)	1	1324	37.14	<0.001	<0.001	0.027 [0.013 - 0.047]	86.42	57.26	57.08	57.44	59.39	c
Environment													
Social Support	1457 (945, 512)	1	1452	487.67	<0.001	<0.001	0.251 [0.210 - 0.289]	53.40	72.17	79.89	64.45	78.45	d
Childhood Maltreatment	1460 (947, 513)	1	1455	381.61	<0.001	<0.001	0.208 [0.169 - 0.244]	53.90	70.84	80.68	61.01	76.23	d

BACC = balanced accuracy, AUROC = area under the receiver operating curve, HC = Healthy Controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, FA = fractional anisotropy, MD = mean diffusivity. Covariates in the statistical models: a= age + sex + dummy scanner + total intracranial volume, b= age + sex + dummy scanner, c= age + sex + dummy site + MDS 1 + MDS 2 + MDS 3, d= age + sex + dummy site

eTable 4. Statistical Analyses of Differences Between Healthy and Depressive Individuals Analyzed Separately for Females and Males

HC versus MDD (male)	**n (HC, MDD)**	***df1***	***df2***	***F***	p_{uncorr}	p_{corr}	$\eta^2_{partial}$	**overlap, %**	**BACC, %**	**sens, %**	**spec, %**	**AUROC, %**	**covariates**
Structural MRI													
VBM	612 (330, 282)	1	606	14.62	<0.001	0.020	0.024 [0.006 - 0.054]	87.85	54.45	56.06	52.84	57.99	a
Cortical and subcortical surface, thickness, volume	611 (328, 283)	1	606	9.94	0.002	1.000	0.016 [0.003 - 0.042]	89.91	55.49	57.62	53.36	57.16	b
Task-based functional MRI													
Face matching task	425 (228, 197)	1	420	8.56	0.004	0.557	0.020 [0.002 - 0.052]	84.89	52.94	46.49	59.39	56.01	b
Resting-state fMRI													
Bivariate connectivity	460 (250, 210)	1	455	18.75	<0.001	0.498	0.040 [0.012 - 0.080]	84.00	58.69	58.80	58.57	60.71	b
Network parameters	460 (250, 210)	1	455	16.11	<0.001	0.125	0.034 [0.009 - 0.077]	84.75	56.80	53.60	60.00	60.34	b
Local Correlation	460 (250, 210)	1	455	23.84	<0.001	0.062	0.050 [0.020 - 0.096]	82.50	59.01	60.40	57.62	62.41	b
ALFF	460 (250, 210)	1	455	15.76	<0.001	0.462	0.033 [0.009 - 0.074]	85.69	55.39	58.40	52.38	60.02	b
fALFF	460 (250, 210)	1	455	18.70	<0.001	0.280	0.039 [0.011 - 0.081]	84.38	57.85	57.60	58.10	61.00	b
Structural connectome													
FA	530 (290, 240)	1	525	6.33	0.012	1.000	0.012 [0.001 - 0.034]	89.38	56.60	58.62	54.58	55.96	b
MD	528 (289, 239)	1	523	12.83	<0.001	0.543	0.024 [0.005 - 0.055]	84.16	57.25	64.71	49.79	57.91	b
Network parameters	533 (292, 241)	1	528	9.05	0.003	1.000	0.017 [0.002 - 0.043]	89.73	55.65	54.45	56.85	56.54	b
Genetics													
Polygenic Risk Score	591 (317, 274)	1	584	25.86	<0.001	<0.001	0.042 [0.017 - 0.079]	83.60	58.49	59.31	57.66	61.36	c
Environment													
Social Support	640 (338, 302)	1	636	178.80	<0.001	<0.001	0.219 [0.158 - 0.276]	58.78	71.83	78.11	65.56	78.08	d
Childhood Maltreatment	640 (339, 301)	1	636	121.61	<0.001	<0.001	0.161 [0.111 - 0.209]	60.12	66.83	78.17	55.48	72.92	d
HC versus MDD (female)													
Structural MRI													
VBM	1132 (596, 536)	1	1126	22.35	<0.001	0.109	0.019 [0.006 - 0.038]	89.07	55.50	53.36	57.65	57.50	a

Cortical and subcortical surface, thickness, volume	1130 (595, 535)	1	1125	14.07	<0.001	0.090	0.012 [0.003 - 0.027]	91.20	53.91	53.61	54.21	55.81	b
Task-based functional MRI													
Face matching task	816 (428, 388)	1	811	14.95	<0.001	0.628	0.018 [0.004 - 0.041]	89.56	55.57	54.44	56.70	56.96	b
Functional connectome													
Bivariate connectivity	876 (452, 424)	1	871	16.22	<0.001	1.000	0.018 [0.005 - 0.041]	88.52	55.29	53.98	56.60	57.35	b
Network parameters	876 (452, 424)	1	871	10.26	0.001	1.000	0.012 [0.001 - 0.029]	91.63	53.81	52.43	55.19	55.39	b
Local Correlation	876 (452, 424)	1	871	17.37	<0.001	0.438	0.020 [0.006 - 0.042]	89.12	55.95	55.53	56.37	57.89	b
ALFF	876 (452, 424)	1	871	15.75	<0.001	0.629	0.018 [0.004 - 0.040]	88.67	55.47	63.05	47.88	57.93	b
fALFF	876 (452, 424)	1	871	19.00	<0.001	0.562	0.021 [0.005 - 0.043]	88.43	55.73	55.09	56.37	57.87	b
Structural connectome													
FA	975 (527, 448)	1	970	10.88	0.001	1.000	0.011 [0.002 - 0.029]	91.51	54.65	52.37	56.92	55.36	b
MD	939 (511, 428)	1	934	15.17	<0.001	0.144	0.016 [0.003 - 0.033]	85.22	52.86	46.38	59.35	55.76	b
Network parameters	975 (527, 448)	1	970	9.33	0.002	1.000	0.010 [0.001 - 0.025]	91.62	54.02	50.66	57.37	55.21	b
Genetics													
Polygenic Risk Score	1053 (549, 504)	1	1046	25.99	<0.001	<0.001	0.024 [0.009 - 0.045]	87.56	57.92	55.92	59.92	58.96	c
Environment													
Social Support	1157 (607, 550)	1	1153	302.90	<0.001	<0.001	0.208 [0.170 - 0.245]	54.29	69.89	79.24	60.55	75.57	d
Childhood Maltreatment	1159 (608, 551)	1	1155	306.47	<0.001	<0.001	0.210 [0.175 - 0.246]	53.33	72.40	80.92	63.88	77.48	d

BACC = balanced accuracy, AUROC = area under the receiver operating curve, HC = Healthy Controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, FA = fractional anisotropy, MD = mean diffusivity. Covariates in the statistical models: a= age + dummy scanner + total intracranial volume, b= age + dummy scanner, c= age + dummy site + MDS 1 + MDS 2 + MDS 3, d= age + dummy site

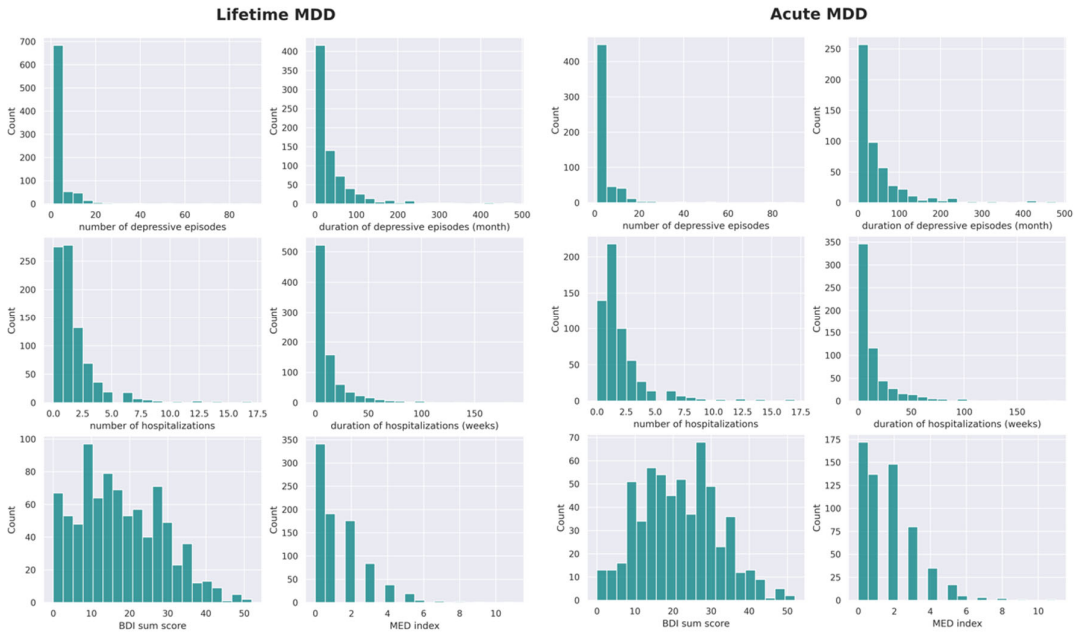
eTable 5. Statistical Analyses of Differences Between Healthy and Depressive Individuals Analyzed Separately for Marburg and Münster

HC versus MDD (Marburg)	**n (HC, MDD)**	***df1***	***df2***	***F***	p_{uncorr}	p_{corr}	$\eta^2_{partial}$	**overlap, %**	**BACC, %**	**sens, %**	**spec, %**	**AUROC, %**	**covariates**
Structural MRI													
VBM	914 (543, 371)	1	909	23.44	<0.001	0.017	0.025 [0.009 - 0.046]	87.02	55.69	54.51	56.87	58.69	a
Cortical and subcortical surface, thickness, volume	914 (541, 373)	1	910	8.99	0.003	0.718	0.010 [0.001 - 0.026]	91.86	54.92	57.30	52.55	56.03	b
Task-based functional MRI													
Face matching task	806 (493, 313)	1	802	13.07	<0.001	0.558	0.016 [0.003 - 0.034]	88.98	55.48	55.38	55.59	57.81	b
Functional connectome													
Bivariate connectivity	870 (524, 346)	1	872	26.57	<0.001	0.014	0.030 [0.013 - 0.057]	85.33	55.78	56.11	55.49	59.40	b
Network parameters	870 (524, 346)	1	872	15.02	<0.001	0.221	0.017 [0.004 - 0.036]	89.29	54.92	53.63	56.07	57.25	b
Local Correlation	870 (524, 346)	1	866	14.76	<0.001	0.496	0.017 [0.005 - 0.039]	89.16	55.22	55.53	54.91	57.63	b
ALFF	870 (524, 346)	1	866	22.36	<0.001	0.018	0.025 [0.010 - 0.051]	85.11	54.71	62.02	47.40	58.28	b
fALFF	870 (524, 346)	1	866	17.70	<0.001	0.470	0.020 [0.004 - 0.041]	88.40	55.56	56.49	54.62	58.32	b
Structural connectome													
FA	778 (473, 305)	1	774	11.15	0.001	1.000	0.014 [0.002 - 0.035]	90.21	55.22	56.66	53.77	57.06	b
MD	798 (491, 307)	1	794	15.60	<0.001	0.121	0.019 [0.005 - 0.042]	88.08	55.43	49.29	61.56	57.78	b
Network parameters	815 (499, 316)	1	811	13.01	<0.001	0.735	0.016 [0.002 - 0.038]	89.37	55.74	52.30	59.18	57.65	b
Genetics													
Polygenic Risk Scores	867 (506, 361)	1	860	46.50	<0.001	<0.001	0.051 [0.028 - 0.086]	81.43	58.95	58.89	59.00	62.99	c
Environment													
Social Support	952 (561, 391)	1	948	293.58	<0.001	<0.001	0.236 [0.185 - 0.285]	55.35	70.86	79.32	62.40	78.57	d
Childhood Maltreatment	950 (561, 389)	1	946	269.91	<0.001	<0.001	0.222 [0.177 - 0.270]	52.88	72.87	82.00	63.75	78.42	d
HC versus MDD (Münster)													
Structural MRI													
VBM	830 (383, 447)	1	825	14.03	<0.001	0.016	0.017 [0.005 - 0.039]	89.69	53.98	50.91	57.05	57.00	a
Cortical and subcortical surface, thickness, volume	827 (382, 445)	1	823	7.84	0.005	1.000	0.009 [0.001 - 0.027]	92.30	53.74	52.88	54.61	55.97	b

Task-based functional MRI													
Face matching task	435 (163, 272)	1	431	15.64	<0.001	0.332	0.035 [0.008 - 0.075]	84.41	56.22	55.83	56.62	60.44	b
Functional connectome													
Bivariate connectivity	466 (178, 288)	1	465	13.01	<0.001	1.000	0.027 [0.005 - 0.060]	85.98	59.11	57.87	59.03	60.04	b
Network parameters	466 (178, 288)	1	465	10.32	0.001	1.000	0.022 [0.003 - 0.058]	87.72	56.75	53.37	60.42	58.59	b
Local Correlation	466 (178, 288)	1	462	18.45	<0.001	0.260	0.038 [0.014 - 0.075]	82.51	58.53	60.11	56.94	60.39	b
ALFF	466 (178, 288)	1	462	29.12	<0.001	0.017	0.059 [0.023 - 0.110]	77.76	58.98	53.37	64.58	62.38	b
fALFF	466 (178, 288)	1	462	15.74	<0.001	0.689	0.033 [0.008 - 0.071]	84.95	57.88	56.74	59.03	60.58	b
Structural connectome													
FA	692 (320, 372)	1	688	10.18	0.001	1.000	0.015 [0.002 - 0.037]	90.52	53.80	52.50	55.11	56.15	b
MD	690 (318, 372)	1	686	14.48	<0.001	0.210	0.021 [0.006 - 0.044]	88.42	56.05	48.11	63.98	58.48	b
Network parameters	693 (320, 373)	1	689	7.91	0.005	1.000	0.011 [0.001 - 0.032]	91.23	54.44	62.50	46.38	56.11	b
Genetics													
Polygenic Risk Scores	752 (343, 409)	1	745	13.53	<0.001	<0.001	0.018 [0.004 - 0.042]	89.41	57.52	56.85	58.19	57.49	c
Environment													
Social Support	845 (384, 461)	1	841	193.68	<0.001	<0.001	0.187 [0.147 - 0.228]	55.62	70.21	78.39	62.04	75.26	d
Childhood Maltreatment	849 (386, 463)	1	845	157.94	<0.001	<0.001	0.157 [0.122 - 0.195]	57.73	69.34	77.98	60.69	73.67	d

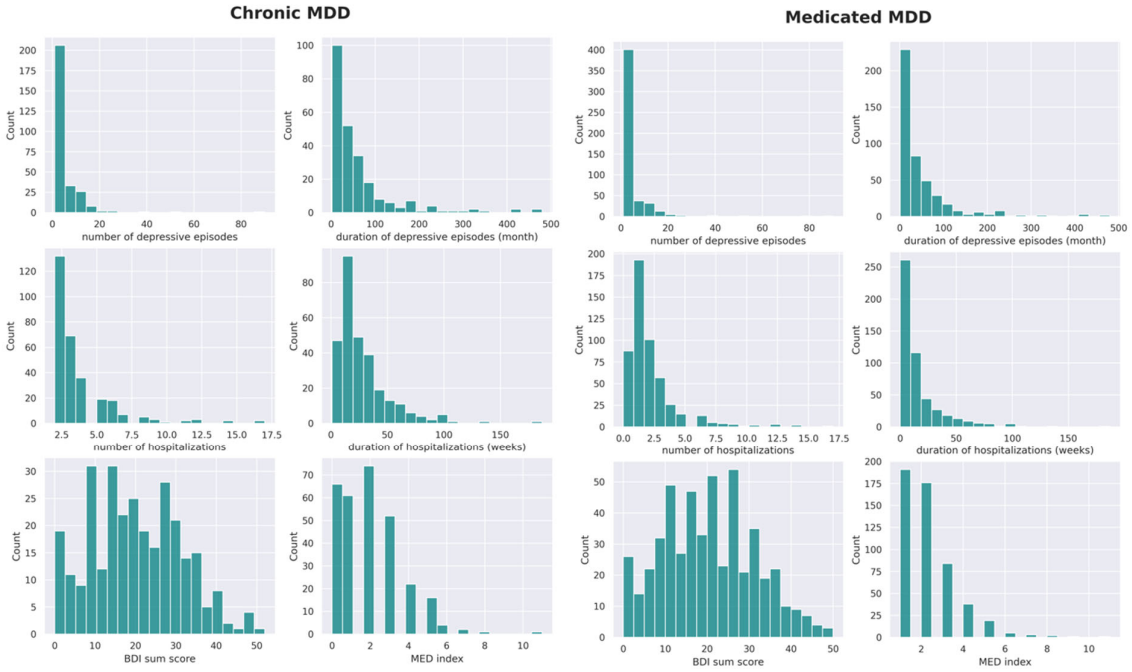
BACC = balanced accuracy, AUROC = area under the receiver operating curve, HC = Healthy Controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, FA = fractional anisotropy, MD = mean diffusivity. Covariates in the statistical models: a= age + sex + total intracranial volume, b= age + sex, c= age + sex + MDS 1 + MDS 2 + MDS 3, d= age + sex

eFigure 1. Depressive Symptom Severity of Lifetime and Acute MDD



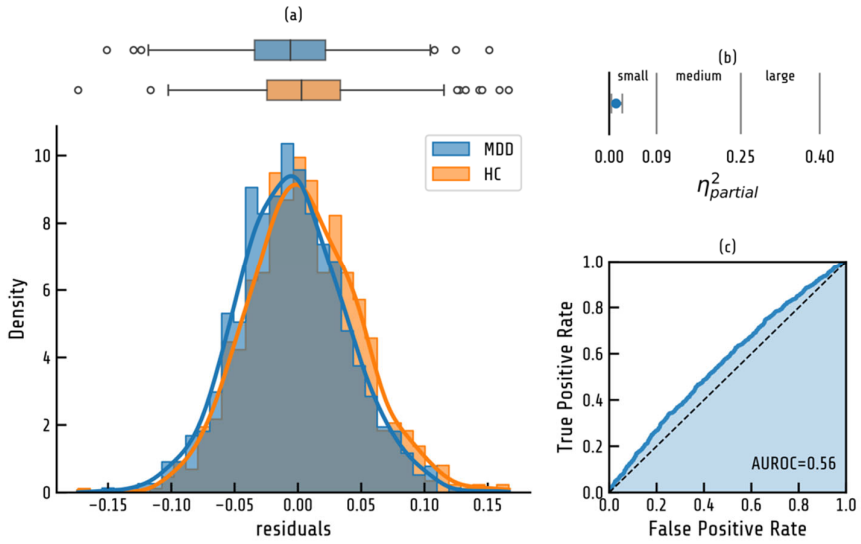
eFigure 1. Histograms of variables describing depressive symptom severity of lifetime and acutely depressed MDD sample.

eFigure 2. Depressive Symptom Severity of Chronically Depressed and Medicated MDD



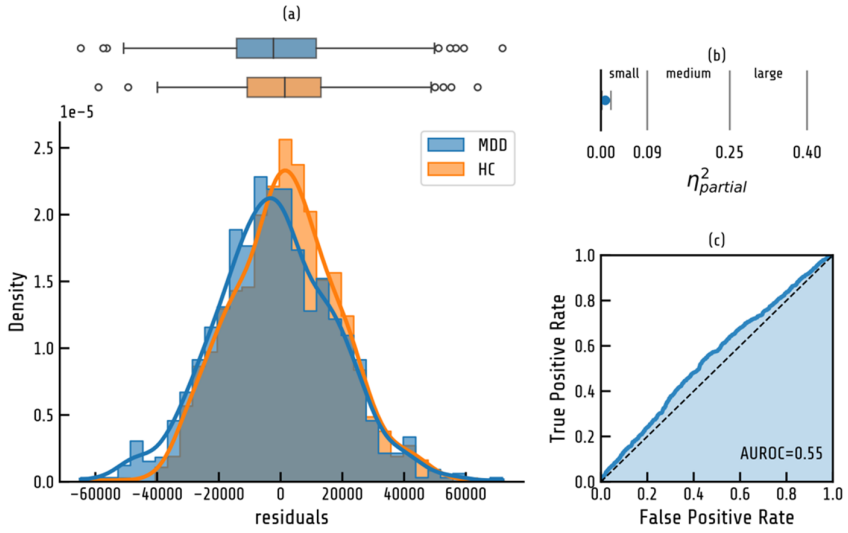
eFigure 2. Histograms of variables describing depressive symptom severity of chronically depressed and medicated MDD sample.

eFigure 3. Predictive Utility—VBM—Healthy Individuals vs Those With Depression



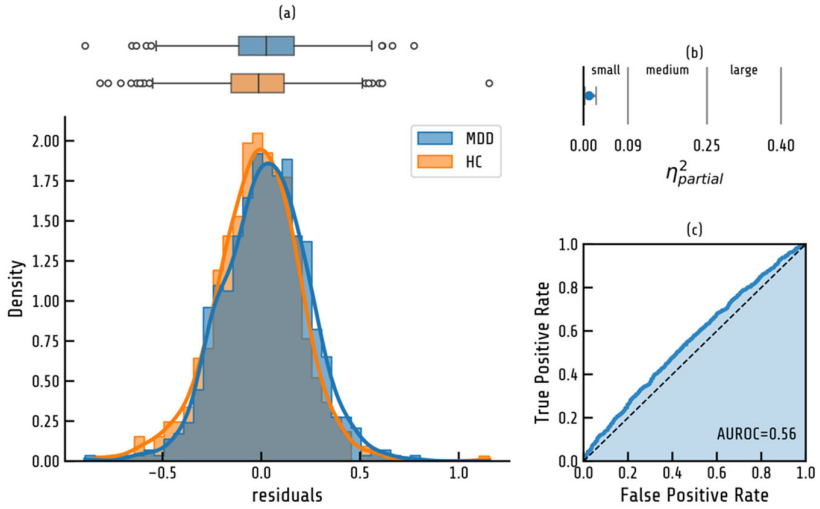
eFigure 3. Distributional overlap, effect size and classification performance for VBM data. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the voxel displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 4. Predictive Utility—Freesurfer—Healthy Individuals vs Those With Depression



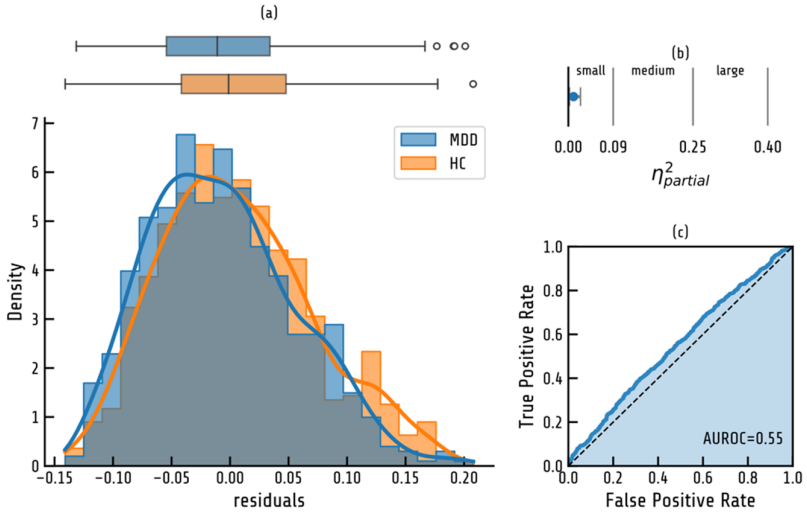
eFigure 4. Distributional overlap, effect size and classification performance for Freesurfer data. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the Freesurfer region displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 5. Predictive Utility—Task fMRI—Healthy Individuals vs Those With Depression



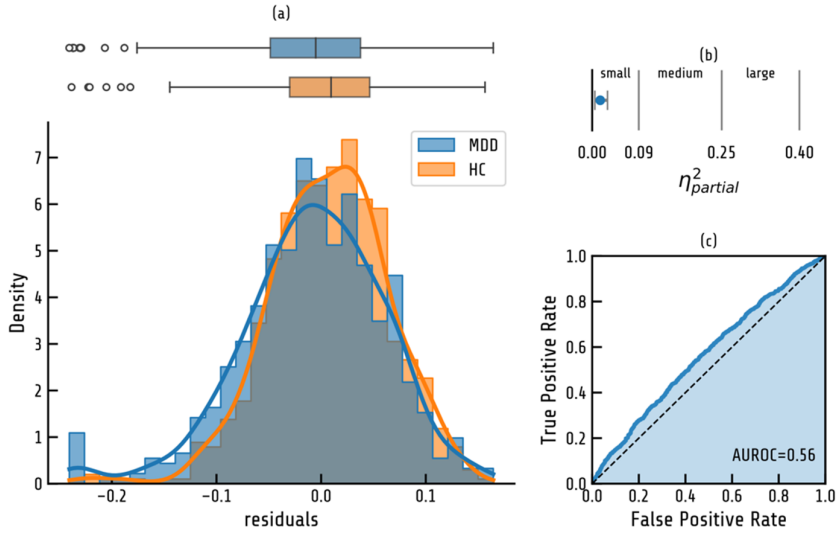
eFigure 5. Distributional overlap, effect size and classification performance for task-based fMRI data. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the voxel displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 6. Predictive Utility—Resting-State fMRI Network Parameters—Healthy Individuals vs Those With Depression



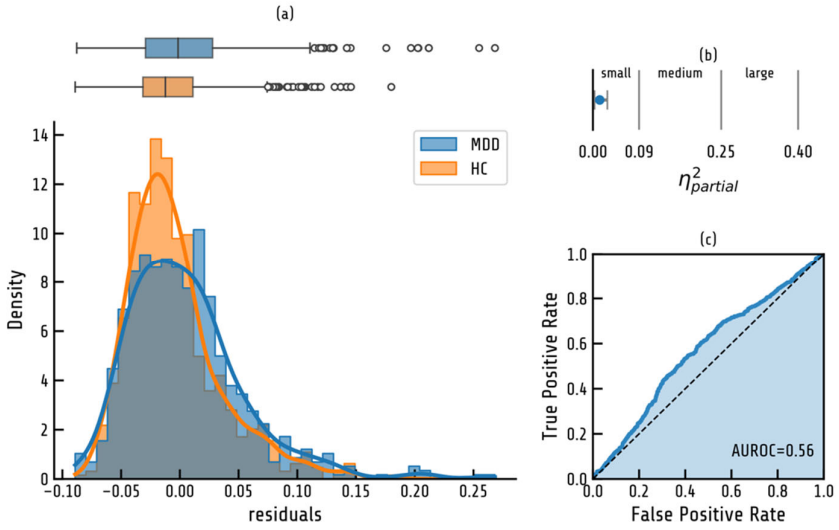
eFigure 6. Distributional overlap, effect size and classification performance for resting state (RS) fMRI network graph metrics. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the network graph parameter displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 7. Predictive Utility—Resting-State fMRI local correlation—Healthy Individuals vs Those With Depression



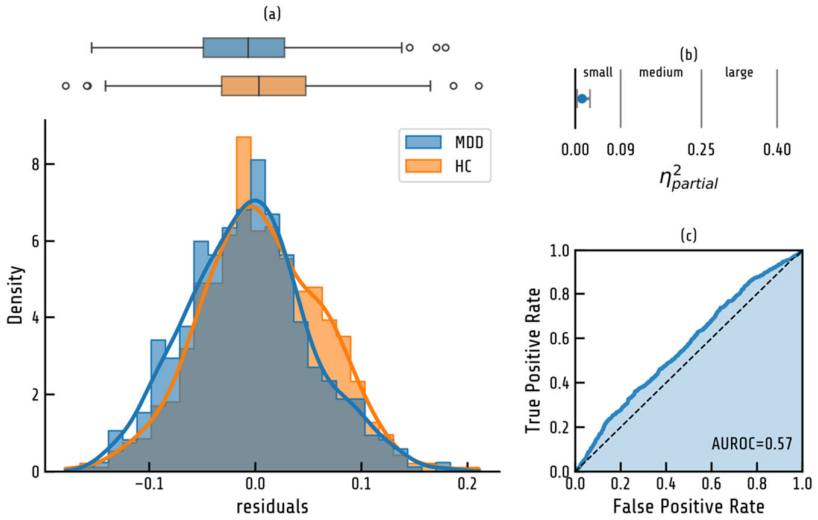
eFigure 7. Distributional overlap, effect size and classification performance for resting state (RS) fMRI local correlation. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the voxel displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 8. Predictive Utility—Resting-State fMRI ALFF—Healthy Individuals vs Those With Depression



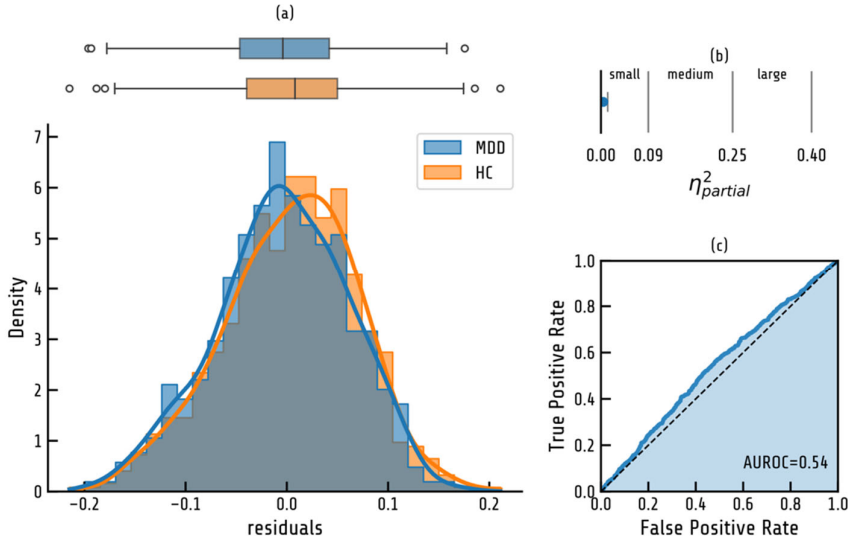
eFigure 8. Distributional overlap, effect size and classification performance for resting state (RS) fMRI ALFF. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the voxel displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 9. Predictive Utility—Resting-State fMRI fALFF—Healthy Individuals vs Those With Depression



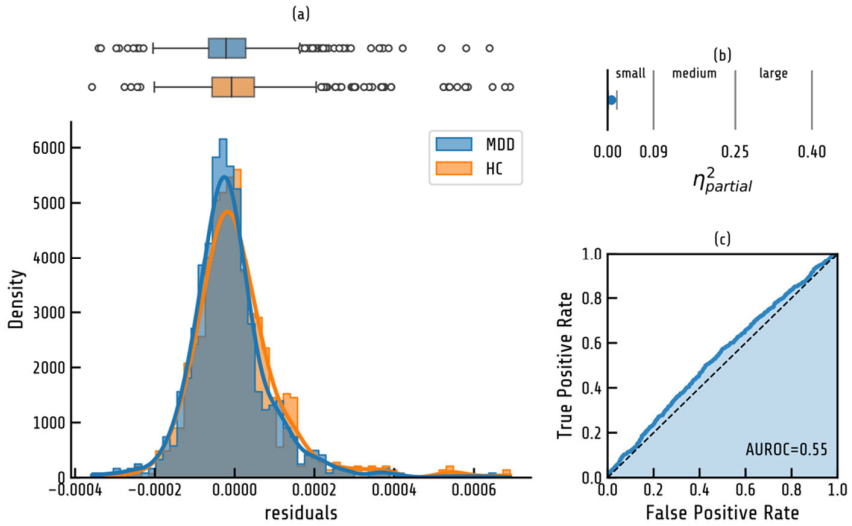
eFigure 9. Distributional overlap, effect size and classification performance for resting state (RS) fMRI fALFF. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the voxel displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 10. Predictive Utility—DTI FA—Healthy Individuals vs Those With Depression



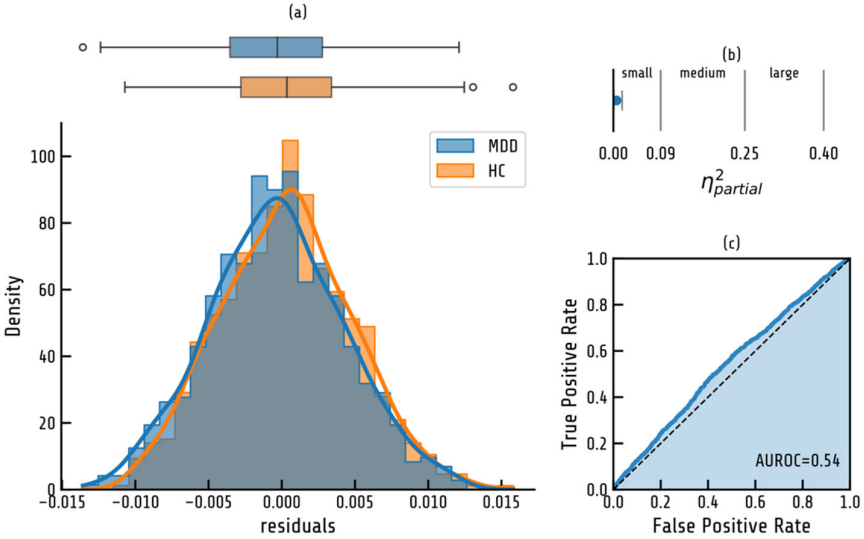
eFigure 10. Distributional overlap, effect size and classification performance for DTI FA. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the connection displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 11. Predictive Utility—DTI MD—Healthy Individuals vs Those With Depression



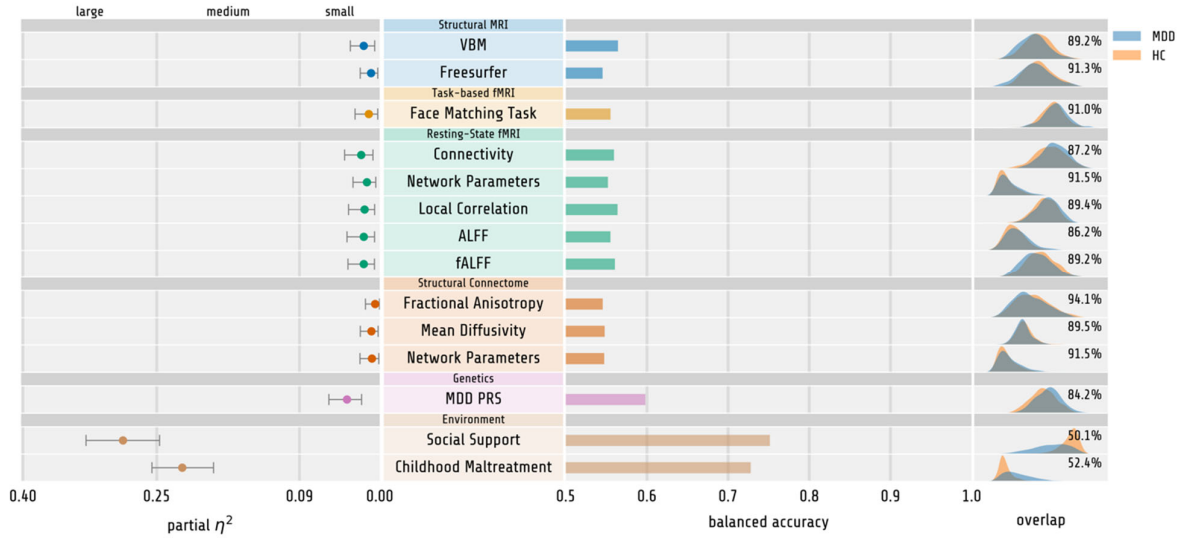
eFigure 11. Distributional overlap, effect size and classification performance for DTI MD. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the connection displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 12. Predictive Utility—DTI Network Parameters—Healthy Individuals vs Those With Depression



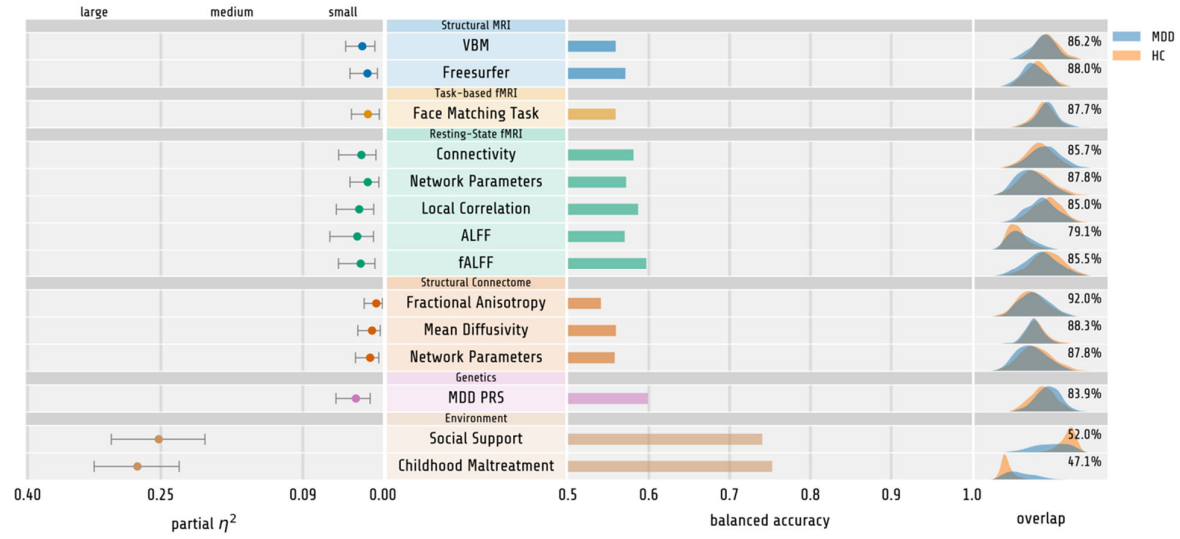
eFigure 12. Distributional overlap, effect size and classification performance for DTI graph metrics. (a) shows a histogram with Gaussian Kernel Density estimation as solid line and boxplot of the confound-corrected values of the network graph parameter displaying the largest effect. (b) partial η^2 ANOVA effect size. Light blue indicates upper bound of bootstrapped 95% confidence interval. (c) shows Receiver-Operating-Characteristic (ROC) Curve for Logistic Regression classification.

eFigure 13. Effect Size and Classification Accuracy—Healthy Individuals vs Those With Acute Depression



eFigure 13. Left: Partial η^2 effect size of single variables displaying the overall largest effect in each modality. Error bars indicate upper and lower bound for bootstrapped confidence intervals for partial η^2. Right: Balanced classification accuracy for all modalities based on Logistic Regression of single variables displaying the largest effect. Kernel density estimation plots of deconfounded values including distributional overlap for healthy and depressive participants are plotted on the right side of the figure. HC = healthy controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, DTI = Diffusion Tensor Imaging, FA = Fractional Anisotropy, MD = Mean Diffusivity, PRS = Polygenic Risk Score.

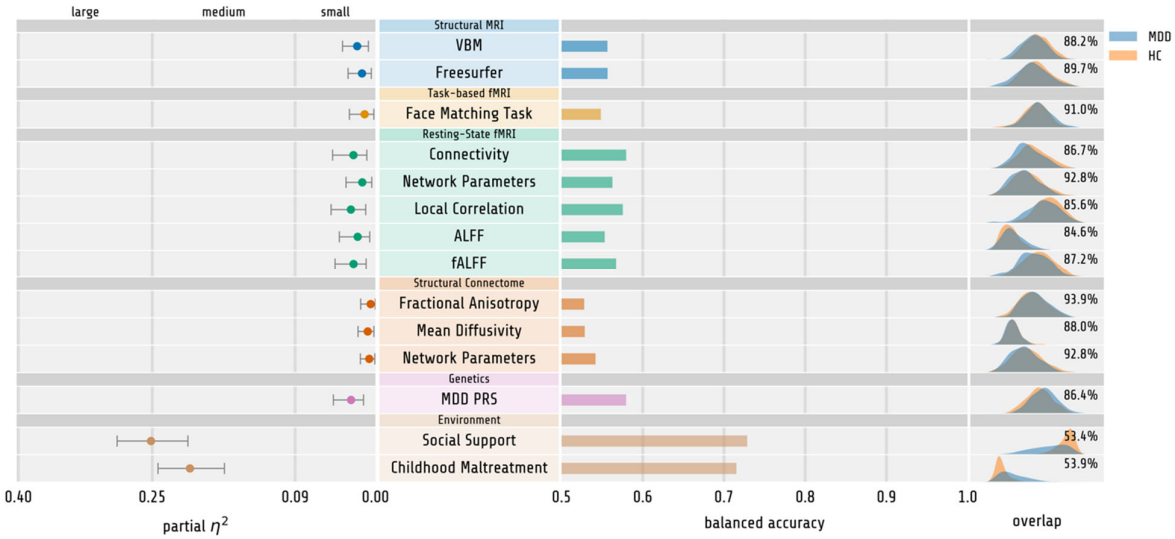
eFigure 14. Effect Size and Classification Accuracy—Healthy Individuals vs Those With Long-term Depression



eFigure 14. Left: Partial η^2 effect size of single variables displaying the overall largest effect in each modality. Error bars indicate upper and lower bound for bootstrapped confidence intervals for partial η^2. Right: Balanced classification accuracy for all modalities based on Logistic Regression of single variables displaying the largest effect. Kernel density estimation plots of deconfounded values including distributional overlap for healthy and depressive participants are plotted on the right

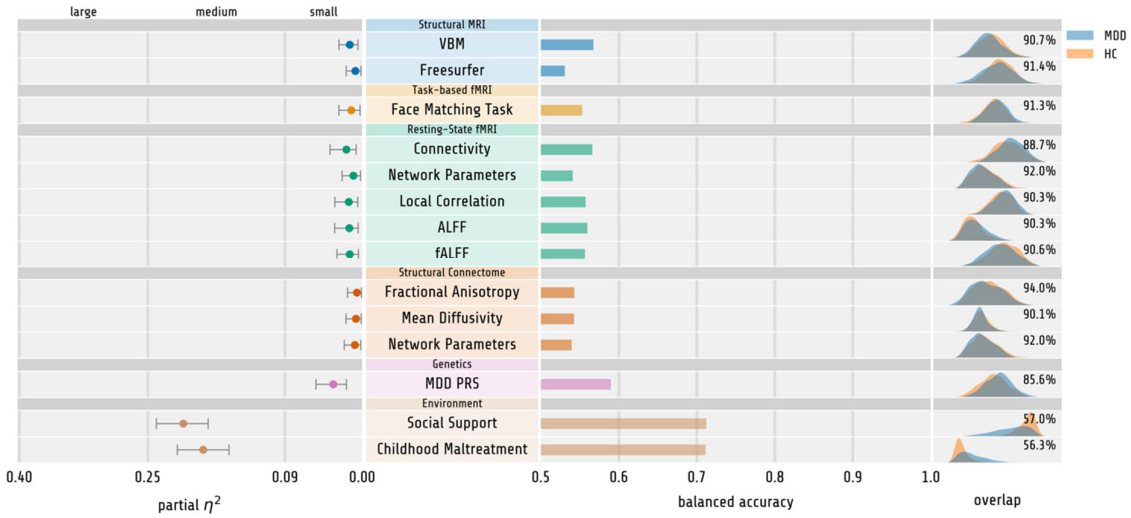
side of the figure. HC = healthy controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, DTI = Diffusion Tensor Imaging, FA = Fractional Anisotropy, MD = Mean Diffusivity, PRS = Polygenic Risk Score.

eFigure 15. Effect Size and Classification Accuracy—Healthy Individuals vs Those With Depression Taking Medication



eFigure 15. Left: Partial η^2 effect size of single variables displaying the overall largest effect in each modality. Error bars indicate upper and lower bound for bootstrapped confidence intervals for partial η^2. Right: Balanced classification accuracy for all modalities based on Logistic Regression of single variables displaying the largest effect. Kernel density estimation plots of deconfounded values including distributional overlap for healthy and depressive participants are plotted on the right side of the figure. HC = healthy controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, DTI = Diffusion Tensor Imaging, FA = Fractional Anisotropy, MD = Mean Diffusivity, PRS = Polygenic Risk Score.

eFigure 16. Effect Size and Classification Accuracy—Matched Healthy Individuals vs Those With Depression



eFigure 16. Left: Partial η^2 effect size of single variables displaying the overall largest effect in each modality. Error bars indicate upper and lower bound for bootstrapped confidence intervals for partial η^2. Right: Balanced classification accuracy for all modalities based on Logistic Regression of single variables displaying the largest effect. Kernel density estimation plots of deconfounded values including distributional overlap for healthy and depressive participants are plotted on the right side of the figure. HC = healthy controls, MDD = Major Depressive Disorder, VBM = Voxel-Based Morphometry, ALFF = Amplitude of Low-Frequency Fluctuations, fALFF = fractional Amplitude of Low-Frequency Fluctuations, DTI = Diffusion Tensor Imaging, FA = Fractional Anisotropy, MD = Mean Diffusivity, PRS = Polygenic Risk Score.

eReferences

1. Wittchen HU, Wunderlich U, Gruschwitz S, Zaudig M. SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. Eine deutschsprachige, erweiterte Bearb. d. amerikanischen Originalversion des SKID I. Published online 1997.

2. Benkert O, Hippius H. *Kompendium Der Psychiatrischen Pharmakotherapie*. Springer, Berlin, Heidelberg; 2021. doi:10.1007/978-3-662-61753-3

3. Redlich R, Dohm K, Grotegerd D, et al. Reward Processing in Unipolar and Bipolar Depression: A Functional MRI Study. *Neuropsychopharmacol*. 2015;40(11):2623-2631. doi:10.1038/npp.2015.110

4. Redlich R, Almeida JR, Grotegerd D, et al. Brain Morphometric Biomarkers Distinguishing Unipolar and Bipolar Depression: A Voxel-Based Morphometry–Pattern Classification Approach. *Jama Psychiat*. 2014;71(11):1222-1230. doi:10.1001/jamapsychiatry.2014.1100

5. Pelin H, Ising M, Stein F, et al. Identification of transdiagnostic psychiatric disorder subtypes using unsupervised learning. *Neuropsychopharmacol*. 2021;46(11):1895-1905. doi:10.1038/s41386-021-01051-0

6. Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*. 2001;29(4):1165. http://www.jstor.org/stable/2674075

7. Ho DE, Imai K, King G, Stuart EA. MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw*. 2011;42(8). doi:10.18637/jss.v042.i08

8. Bernstein DP, Fink L, Handelsman L, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am J Psychiat*. 1994;151(8):1132-1136. doi:10.1176/ajp.151.8.1132

9. Fydrich T, Sommer G, Tydecks S, Brähler E. Fragebogen zur sozialen Unterstützung (F-SozU): Normierung der Kurzform (K-14). *Zeitschrift für Medizinische Psychologie*. 2009;18:43.

10. Andlauer TFM, Buck D, Antony G, et al. Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. *Sci Adv*. 2016;2(6):e1501678. doi:10.1126/sciadv.1501678

11. Meller T, Schmitt S, Stein F, et al. Associations of schizophrenia risk genes ZNF804A and CACNA1C with schizotypy and modulation of attention in healthy subjects. *Schizophr Res*. 2019;208:67-75. doi:10.1016/j.schres.2019.04.018

12. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1):7. doi:10.1186/s13742-015-0047-8

13. Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *Plos Genet*. 2009;5(6):e1000529. doi:10.1371/journal.pgen.1000529

14. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*. 2012;44(8):955-959. doi:10.1038/ng.2354

15. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019;10(1):1776. doi:10.1038/s41467-019-09718-5

16. Howard DM, Adams MJ, Clarke TK, 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(3):343-352. doi:10.1038/s41593-018-0326-7

17. Vogelbacher C, Möbius TWD, Sommer J, et al. The Marburg-Münster Affective Disorders Cohort Study (MACS): A quality assurance protocol for MR neuroimaging data. *Neuroimage*. 2018;172:450-460. doi:10.1016/j.neuroimage.2018.01.079

18. Gaser C, Kurth F. *Computational Anatomy Toolbox CAT12*. http://www.neuro.uni-jena.de/cat/

19. Fischl B. FreeSurfer. *Neuroimage*. 2012;62(2):774-781. doi:10.1016/j.neuroimage.2012.01.021

20. Consortium E. Structural image processing protocols. http://enigma.ini.usc.edu/protocols/imaging-protocols/

21. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*. 2012;2(3):125-141. doi:10.1089/brain.2012.0073

22. Muschelli J, Nebel MB, Caffo BS, Barber AD, Pekar JJ, Mostofsky SH. Reduction of motion-related artifacts in resting state fMRI using aCompCor. *Neuroimage*. 2014;96:22-35. doi:10.1016/j.neuroimage.2014.03.028

23. Schaefer A, Kong R, Gordon EM, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex*. 2018;28(9):3095-3114. doi:10.1093/cercor/bhx179

24. Yang H, Long XY, Yang Y, et al. Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI. *Neuroimage*. 2007;36(1):144-152. doi:10.1016/j.neuroimage.2007.01.054

25. Zou QH, Zhu CZ, Yang Y, et al. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J Neurosci Meth*. 2008;172(1):137-141. doi:10.1016/j.jneumeth.2008.04.012

26. Deshpande G, LaConte S, Peltier S, Hu X. Integrated local correlation: A new measure of local coherence in fMRI data. *Hum Brain Mapp*. 2009;30(1):13-23. doi:10.1002/hbm.20482

27. Hariri AR, Tessitore A, Mattay VS, Fera F, Weinberger DR. The Amygdala Response to Emotional Stimuli: A Comparison of Faces and Scenes. *Neuroimage*. 2002;17(1):317-323. doi:10.1006/nimg.2002.1179

28. Kessler R, Schmitt S, Sauder T, et al. Long-Term Neuroanatomical Consequences of Childhood Maltreatment: Reduced Amygdala Inhibition by Medial Prefrontal Cortex. *Frontiers Syst Neurosci*. 2020;14:28. doi:10.3389/fnsys.2020.00028

29. Repple J, Mauritz M, Meinert S, et al. Severity of current depression and remission status are associated with structural connectome alterations in major depressive disorder. *Mol Psychiatr*. 2020;25(7):1550-1558. doi:10.1038/s41380-019-0603-1

30. Heuvel MP van den, Sporns O, Collin G, et al. Abnormal Rich Club Organization and Functional Brain Dynamics in Schizophrenia. *Jama Psychiat*. 2013;70(8):783-792. doi:10.1001/jamapsychiatry.2013.1328

31. Andersson JLR, Skare S. A Model-Based Method for Retrospective Correction of Geometric Distortions in Diffusion-Weighted EPI. *Neuroimage*. 2002;16(1):177-199. doi:10.1006/nimg.2001.1039

32. Chang L, Walker L, Pierpaoli C. Informed RESTORE: A method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. *Magnet Reson Med*. 2012;68(5):1654-1663. doi:10.1002/mrm.24173

33. Chang L, Jones DK, Pierpaoli C. RESTORE: Robust estimation of tensors by outlier rejection. *Magnet Reson Med*. 2005;53(5):1088-1095. doi:10.1002/mrm.20426

34. Lange SC de, Heuvel MP van den. Structural and functional connectivity reconstruction with CATO - A Connectivity Analysis TOolbox. *Biorxiv*. Published online 2021:2021.05.31.446012. doi:10.1101/2021.05.31.446012

35. Cammoun L, Gigandet X, Meskaldji D, et al. Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J Neurosci Meth*. 2012;203(2):386-397. doi:10.1016/j.jneumeth.2011.09.031

36. Hagmann P, Cammoun L, Gigandet X, et al. Mapping the Structural Core of Human Cerebral Cortex. *Plos Biol*. 2008;6(7):e159. doi:10.1371/journal.pbio.0060159

37. Lange SC de, Scholtens LH, Initiative ADN, et al. Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav*. 2019;3(9):988-998. doi:10.1038/s41562-019-0659-6

38. Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. *Proceedings of the 7th Python in Science Conference*. Published online 2008:11-15.

39. Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*. 2010;52(3):1059-1069. doi:10.1016/j.neuroimage.2009.10.003