Received: 21 June 2019 | Revised: 8 September 2019 | Accepted: 9 September 2019

DOI: 10.1002/hbm.24802

RESEARCH ARTICLE

WILEY

Multimodal fusion of structural and functional brain imaging in depression using linked independent component analysis

Luigi A. Maglanoc[1,2] | Tobias Kaufmann[2] | Rune Jonassen[3] | Eva Hilland[1,4] | Dani Beck[2,5] | Nils Inge Landrø[1] | Lars T. Westlye[2,5]

[1]Clinical Neuroscience Research Group, Department of Psychology, University of Oslo, Oslo, Norway

[2]NORMENT, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[3]Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

[4]Division of Psychiatry, Diakonhjemmet Hospital, Oslo, Norway

[5]Department of Psychology, University of Oslo, Oslo, Norway

Correspondence
Luigi A. Maglanoc and Lars Westlye, Department of Psychology, University of Oslo, Pb. 1094, Blindern, 0317 Oslo, Norway.
Email: luigi_maglanoc@hotmail.com; l.t.westlye@psykologi.uio.no

Funding information
Helse Sør-Øst RHF, Grant/Award Numbers: 2014097, 2015052, 2015073; Norges Forskningsråd, Grant/Award Numbers: 175387/V50, 229135, 249795; Department of Psychology, University of Oslo; European Research Council, Grant/Award Number: ERC StG 802998

Abstract

Previous structural and functional neuroimaging studies have implicated distributed brain regions and networks in depression. However, there are no robust imaging biomarkers that are specific to depression, which may be due to clinical heterogeneity and neurobiological complexity. A dimensional approach and fusion of imaging modalities may yield a more coherent view of the neuronal correlates of depression. We used linked independent component analysis to fuse cortical macrostructure (thickness, area, gray matter density), white matter diffusion properties and resting-state functional magnetic resonance imaging default mode network amplitude in patients with a history of depression (n = 170) and controls (n = 71). We used univariate and machine learning approaches to assess the relationship between age, sex, case–control status, and symptom loads for depression and anxiety with the resulting brain components. Univariate analyses revealed strong associations between age and sex with mainly global but also regional specific brain components, with varying degrees of multimodal involvement. In contrast, there were no significant associations with case–control status, nor symptom loads for depression and anxiety with the brain components, nor any interaction effects with age and sex. Machine learning revealed low model performance for classifying patients from controls and predicting symptom loads for depression and anxiety, but high age prediction accuracy. Multimodal fusion of brain imaging data alone may not be sufficient for dissecting the clinical and neurobiological heterogeneity of depression. Precise clinical stratification and methods for brain phenotyping at the individual level based on large training samples may be needed to parse the neuroanatomy of depression.

KEYWORDS
depression, heterogeneity, linked independent component analysis, machine learning, multimodal MRI

1 | INTRODUCTION

With an estimated prevalence of 4.4%, depression affects more than 300 million worldwide (World Health Organization, 2017) and is a substantial contributor to disability and health loss (Friedrich, 2017). Identifying useful imaging based and other biomarkers to aid detection of individuals at risk for depression and facilitating individualized treatment is a global aim (Cuthbert & Insel, 2012; Insel, 2014, 2015).

A host of studies across a range of neuroimaging modalities have implicated various brain regions and networks in depression. Meta-analyses of structural magnetic resonance imaging (MRI) studies have suggested thinner orbitofrontal and anterior cingulate cortex in patients with depression compared to healthy controls (Lai, 2013; Schmaal et al., 2017; Suh et al., 2019). A large-scale meta-analysis comprising 2,148 patients and 7,957 controls from 20 different cohorts reported slightly smaller hippocampal volumes in patients with depression compared to controls (Schmaal et al., 2016), but the overall pattern of results suggested substantial heterogeneity and otherwise striking similarity across groups for all other investigated subcortical structures (Fried & Kievit, 2016). A meta-analysis of diffusion tensor imaging (DTI) studies including 641 patients and 581 healthy controls reported fractional anisotropy (FA) reductions in the genu of the corpus callosum and the anterior limb of the internal capsule (Chen et al., 2016), implicating interhemispheric and frontal–striatal–thalamic connections among the neuronal correlates of depression. Supporting the relevance of brain connectivity in mood disorders, resting-state fMRI studies have reported aberrant connectivity within the default mode network (DMN) in patients with depression compared to healthy controls (Kaiser, Andrews-Hanna, Wager, & Pizzagalli, 2015; Mulders, van Eijndhoven, Schene, Beckmann, & Tendolkar, 2015; Yan et al., 2019).

However, despite meta-analytical evidence suggesting brain aberrations in large groups of patients with depression, the reported effect sizes are small and the direct clinical utility is unclear (Müller et al., 2017; Paulus & Thompson, 2019). One explanation for the lack of robust imaging-based markers in depression may be that previous studies either have focused on a single imaging modality or have analyzed different imaging modalities along separate pipelines and thus failed to model the common variance across features. In contrast, linked independent component analysis (LICA: Groves, Beckmann, Smith, & Woolrich, 2011; Groves et al., 2012) offers an integrated approach by fusing different structural and functional imaging modalities (Groves et al., 2012). LICA identifies modes of variation across modalities and disentangles independent sources of variation that may account both for large and small parts of the total variance that may otherwise be overlooked by conventional approaches. By decomposing the imaging data into a set of independent components, LICA enables an integrated perspective that may improve clinical sensitivity compared to unimodal analyses (Alnæs et al., 2018; Doan, Engvig, Persson, et al., 2017; Francx et al., 2016; Wu et al., 2019).

Apart from the predominantly unimodal approaches in previous imaging studies, large individual differences and heterogeneity in the configuration and load of depressive symptoms represent other factors that could explain the lack of robust imaging markers. Symptom-based approaches have revealed more than 1,000 unique symptom profiles among 3,703 depressed outpatients based on only 12 questionnaire items (Fried & Nesse, 2015), suggesting large heterogeneity. Additionally, depression is highly comorbid with anxiety, with reported rates exceeding 50% (Johansson, Carlbring, Heedman, Paxling, & Andersson, 2013; Lamers et al., 2011). Furthermore, depression can be conceptualized along a continuum including individuals of the general, healthy population that may experience transient symptoms to varying degrees, and thus warrants a dimensional approach.

The main aim of the current study was to determine whether fusion of neuroimaging modalities would capture modes of brain variations which discriminate between patients with a history of depression (n = 170) and healthy controls with no history of depression (n = 71), and which are sensitive to current symptoms of depression and anxiety across groups. To this end, we used LICA to combine measures of cortical macrostructure (cortical surface area and thickness, and gray matter density [GMD]), white matter diffusion properties (DTI-based FA, mean diffusivity [MD], and radial diffusivity [RD]), and resting-state fMRI DMN amplitude.

There is evidence of sex and age differences in the prevalence and clinical characteristics of depression, including lower age at onset of first major depressive episode in women compared to men (Marcus et al., 2005), and longer duration of illness and different symptoms in older compared to younger patients (Husain et al., 2005), which may reflect differential neuronal correlates. Therefore, we tested for main effects of age and sex and their interactions with the resulting brain components' subject weights on group and symptoms.

In addition, we assessed the overall clinical sensitivity of all measures combined using machine learning to classify patients and controls and to predict symptom loads for depression and anxiety, which we compared with age prediction. Based on the above reviewed studies and current models, we anticipated (a) that brain variance related to depression would be captured in components primarily reflecting the previously extended functional neuroanatomy of depression, including limbic and frontotemporal networks and their connections. Irrespective of having a history of depression, we hypothesized (b) several strong age and sex differences, reflecting well-documented age, and sex-related variance in brain structure, including global thickness and volume reductions with increasing age, and larger brain volume and surface area in men compared to women. To the extent that having a history of depression interacts with sex- and age-related processes in the brain, we hypothesized (c) interactions between the age-related trajectories and sex differences identified above with case–control status or symptoms of depression. To increase robustness and generalizability, we corrected for multiple comparisons across all univariate analyses and performed cross-validation and robust model evaluation in the machine learning analyses.

2 | MATERIALS AND METHODS

2.1 | Sample

Patients (n = 194) were primarily recruited from outpatient clinics, while healthy controls (n = 78) were recruited through posters, newspaper advertisements, and social media. The patient group was drawn from two related clinical trials (ClinicalTrials.gov ID NCT0265862 and NCT02931487). All participants were evaluated with the Mini International Neuropsychiatric Interview (M.I.N.I 6.0: Sheehan et al., 1998). Exclusion criteria for all participants were MRI contraindications and a

self-reported history of neurological disorders. The study was approved by the Regional Ethical Committee of South-Eastern Norway, and we obtained a signed informed consent from all the participants. Symptom loads for depression and anxiety were evaluated using the Becks Depression Inventory (BDI-II; Beck, 1996) and the Becks Anxiety Inventory (BAI; Beck & Steer, 1993), respectively. The demographics for the final sample (after exclusions, see below) are shown in Table 1. The range of symptom load for depression and anxiety for the control group was from 0 to 20 and 0 to 19, respectively, while the range for the patient group was from 0 to 51 and 0 to 45, respectively (see Figure 1 for the distributions).

TABLE 1 Demographics of the final sample. Six patients were missing information about (ISCED level, two controls and two patients were missing AUDIT scores, and two controls and four patients were missing DUDIT scores. *p* denotes the *p*-value from group comparisons using chi-square test for sex, handedness, history of additional disorders, and current SSRI medication status while we used Mann–Whitney *U* tests for the rest. Depression severity was based on BDI-II sum score criteria: minimal (0–13), mild (14–19), moderate (20–28), and severe (29–63)

	Controls (*n* = 70)	Patients (*n* = 171)	*p*
Sex (female, %)	46 (66)	120 (70)	.599
Age (mean, *SD*)	41.8 (13.1)	38.7 (13.3)	.092
Education level ISCED (mean)	6.0 1.0)	5.9 (1.2)	.932
Depression symptoms			
BDI-II (mean, *SD*)	1.6 (3.0)	11.6 (10.4)	<.001
Minimal depression severity (*N*)	68	110	N/A
Mild depression severity (*N*)	0	28	N/A
Moderate depression severity (*N*)	1	16	N/A
Severe depression severity (*N*)	0	17	N/A
Anxiety symptoms			
BAI (mean, *SD*)	1.7 (2.8)	8.1 (8.1)	<.001
Other			
AUDIT (mean, *SD*)	4.8 (3.3)	6.3 (5.0)	.127
DUDIT (mean, *SD*)	0.5 (1.9)	0.8 (2.5)	.123
Left handedness (*N*)	7	6	.087
History of anxiety disorder (*N*)	1	50	<.001
History of (hypo)mania (*N*)	0	23	<.001
History of other Axis-I disorders (*N*)	0	23	<.001
No. major depressive episodes (mean, *SD*)	0	4.4 (5.9)	<.001
Currently medicated (SSRI, *N*)	0	52	<.001

Abbreviations: AUDIT, Alcohol Use Disorders Identification Test; BDI-II, Becks Depression Inventory; DUDIT, Drug Use Disorder Identification Test; ISCED, International Standard Classification of Education; SSRI, selective serotonin reuptake inhibitor.

2.2 | Image acquisition

MRI data were obtained on a 3 T Philips Ingenia scanner (Phillips Healthcare) at the Oslo University Hospital using a 32-channel head coil. The same protocol was used for all participants, but there was a change in the phase-encoding direction during the course of study recruitment and data collection which affected the T1-weighted data for four controls and 95 patients, and resting-state fMRI data for 64 patients.

T1-weighted data were collected for 74 controls and 194 patients using a 3D turbo field echo scan with SENSE using the following parameters: acceleration factor = 2; repetition time (TR)/echo time (TE)/flip angle (FA): 3,000 ms/3.61 ms/8°; scan duration: 3 min 16 s, 1 mm isotropic voxels.

Diffusion weighted data were collected for 72 controls and 184 patients using a dual spin echo, single-shot EPI sequence with the following parameters was used: TR/TE = 7,200/86.5 ms, FOV = 224 × 224 mm^2, 112 × 112 matrix, 2.0 mm isotropic voxels; 32 volumes with noncollinear directions (b = 1,000 s/mm^2). Additionally, we acquired two b = 0 volumes with opposite phase polarity (blip up/down volumes).

Resting-state fMRI data were collected using a T2* weighted single-shot gradient echo EPI sequence was acquired for 72 controls and 178 patients with the following parameters: TR/TE/FA = 2,500 ms/30 ms/80°; 3.00 mm isotropic voxels; 45 slices, 200 volumes; scan time ≈ 8.5 min. Participants were instructed to have their eyes open, and refrain from falling asleep.

2.3 | Structural MRI preprocessing

Vertex-wise cortical thickness and surface area measures (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) were estimated based on the T1-weighted scans using FreeSurfer (http://surfer.nmr.mgh.harvard.edu) (Fischl et al., 2002). Details are described elsewhere (Dale et al., 1999; Fischl et al., 1999) but in short, after gray/white boundary and pial reconstruction, cortical thickness was defined as the shortest distance between the surfaces vertex-wise (Dale et al., 1999), before resampling to the Freesurfer common template (fsaverage, 10,242 vertices; Fischl et al., 1999). The vertex-wise expansion or compression was used to calculate vertex-wise maps of arealization. None of the thickness nor surface area data for healthy control (n = 74) nor patients (n = 194) were excluded after visual QC.

2.4 | Voxel-based morphometry

GMD maps were created based on voxel-based morphometry (VBM) using the computational anatomy toolbox (CAT12: http://www.neuro.uni-jena.de/cat/) within SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). This involved brain extraction, gray matter segmentation, and then registration to MNI152 standard space. The resulting images were averaged and flipped along the x axis to create a left–right symmetric, study-specific gray matter template. The modulated gray matter maps were smoothed with a sigma of 4 mm (FWHM = 9.4 mm).

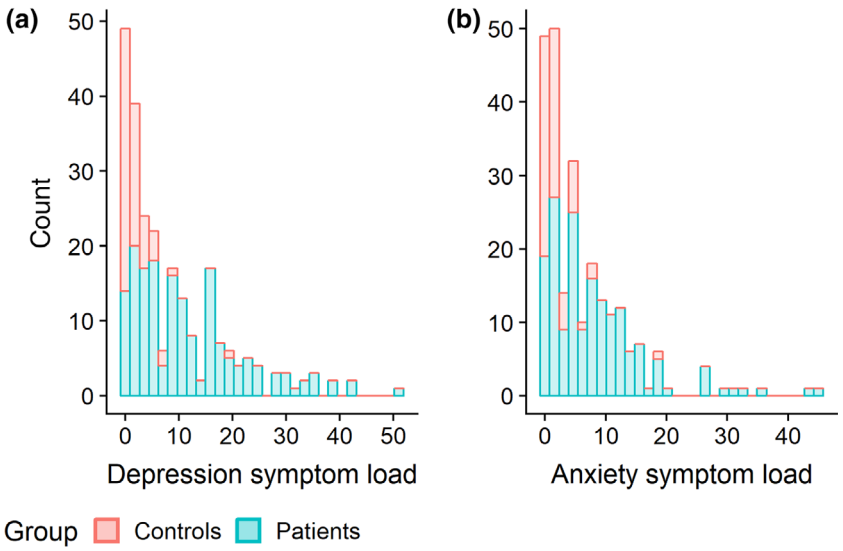

FIGURE 1 Histogram of symptom loads of (a) depression based on Becks Depression Inventory (BDI-II) and (b) anxiety based on Becks Anxiety Inventory (BAI) [Color figure can be viewed at wileyonlinelibrary.com]

None of the GMD data for healthy control (n = 74) nor patients (n = 194) were excluded after visual QC.

2.5 | DTI preprocessing

Processing steps included correction for motion and geometrical distortions based on the two b = 0 volumes and eddy currents by using FSL *topup* (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TOPUP) and *eddy* (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy). We also used *eddy* to automatically identify and replace slices with signal loss (see Figure S1, Supporting Information) within an integrated framework using Gaussian process (Andersson & Sotiropoulos, 2016), which substantially improved the temporal signal-to-noise ratio (tSNR: Roalf et al., 2016) (t = 24.139, p < .001, Cohen's d = 2.13). We fitted a diffusion tensor model using dtifit in FSL to generate maps of FA, MD, and RD. Based on manual QC, we excluded three subjects due to insufficient brain coverage, and three subjects due to poor data quality. One additional subject was flagged with a tSNR of >2 *SD* lower than the mean and discarded after additional manual QC. This yielded a total number of DTI scans of 71 healthy controls and 178 patients.

2.6 | Resting-state fMRI preprocessing

Resting-state fMRI data were processed using the FSL's FMRI Expert Analysis Tool. This included coregistration with T1 images, brain extraction, motion correction (MCFLIRT: Jenkinson, Bannister, Brady, & Smith, 2002), spatial smoothing (FWHM = 6 mm), high-pass filtering (100 s), standard space registration (MNI-152) with FLIRT, and single-session independent component analysis (ICA; MELODIC). Automatic classification and regression of noise components was done using ICA-based Xnoiseifier (FIX: Griffanti et al., 2014; Salimi-Khorshidi et al., 2014), with a threshold of 60. FIX substantially improved tSNR (t = 20.89, p < .001, Cohen's d = 1.95), and no fMRI scans from healthy controls (n = 72) nor from patients (n = 178) were excluded. Group-level ICA with model order fixed at 40 was performed on a balanced subset of healthy controls and patients (N = 72 from each group), which has been used in a previous study (Maglanoc et al., 2019). Dual regression (Nickerson, Smith, Öngür, & Beckmann, 2017) was used to estimate spatial maps and corresponding time series of all components. We then identified an IC representing the canonical DMN (Figure S2, Supporting Information) and used the individual DMN spatial maps from dual regression in multimodal decomposition using LICA.

2.7 | Linked independent component analysis

We used FMRIB's LICA (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLICA) to perform data-driven multimodal fusion, which evaluates shared inter-subject variations across the brain imaging measures (Groves et al., 2011, 2012). LICA is a Bayesian extension of ICA (Hyvarinen, 1999; Jutten & Herault, 1991), providing a factorization over participants rather than time (Groves et al., 2011, 2012). This produces spatial maps based on the commonalities across features (e.g., GMD, DTI measures, DMN maps) and subjects, and corresponding subject weights (i.e., the degree to which a subject contributes to an LICA component). An advantage of this method is its ability to combine modalities with different numbers of spatial dimensions or features by applying an ICA decomposition on each modality while at the same time accounting for the spatial correlation of each modality (Groves et al., 2011, 2012). We included complete data from 70 patients and 171 controls in the decomposition. We chose a relatively low model order of 40 based on previous recommendation of estimating robust components (Wolfers et al., 2017), and the biological meaningfulness of the spatial maps. For transparency and comparison, we also performed similar analysis using a higher dimensionality (80, more details in Supporting Information). For both model orders, we discarded components which were highly driven by one subject (threshold: >20%) yielding a total of 40 and 67 components, respectively (Figure S3, Supporting Information). One component was strongly associated with phase encoding direction (IC4 in both decompositions, t = 33.07

and t = 32.47, p < .001, respectively, Figures S4 and S5, Supporting Information) but not removed from the analyses because of the biologically meaningful spatial patterns.

2.8 | Statistical analysis

Statistical analyses were performed in R version 3.5.1 (R Core Team, 2018) and MATLAB 2014A (The MathWorks). We used linear models to test for main effects of clinical characteristics (case–control status, total symptom severity for depression and anxiety), age, and sex on each LICA subject weight with each IC as the dependent variable. In additional models, we tested for interactions between age or sex and clinical characteristics (case–control status, total symptom severity for depression and anxiety) on each IC. For the analyses involving symptoms, one healthy control was removed due to missing data. We included phase encoding direction as an additional covariate in all the univariate analyses, and we controlled the false discovery rate (FDR) across tests using p.adjust in R.

2.9 | Machine learning approach

For group classification, we submitted all LICA subject weights to shrinkage discriminant analysis (Ahdesmäki & Strimmer, 2010) in the R-package "sda" (http://www.strimmerlab.org/software/sda/). For the main analyses, we used the residuals of each component's subject weight after regressing out age and sex, and additionally, phase encoding for IC4. As a supplemental analysis, we used the residuals of the subject weights after regressing out age, sex, and phase encoding direction from all the ICs. For robustness and to reduce overfitting, we performed cross-validation with 10 folds across 100 iterations. We calculated area under the receiver-operating curve (AUC) as our main measure of model performance using the R-package "pROC" (Robin et al., 2011). We additionally estimated accuracy (i.e., proportion of correct classification), sensitivity (i.e., ability to correctly detect cases) and specificity (i.e., ability to correctly detect controls) using a probability threshold of .65 to partly account for the imbalanced group size. The relative feature importance was determined by calculating correlation-adjusted t-scores (CAT scores) between the group centroids and the pooled mean (Ahdesmäki & Strimmer, 2010). We determined statistical significance based on AUC using permutation-based testing across 10,000 iterations. Several supplemental analyses were conducted to account for imbalanced group sizes and depression severity in the group classification, as well as SSRI medication use (see the Methods section, Supporting Information).

We used the same framework to predict depression and anxiety total symptom severity, but by implementing shrinkage linear estimation (Schäfer & Strimmer, 2005) in the R-package "care" (http://strimmerlab.org/software/care). Here, we computed root mean squared error (RMSE) between the raw and predicted scores as our main measure of model performance, but also mean absolute error (MAE), Spearman's ρ, and R^2. In this case, the relative feature importance was determined by computing the mean correlation-adjusted marginal correlation (CAR) scores, which are the correlations between the response and the Mahalanobis-decorrelated predictors (Zuber &

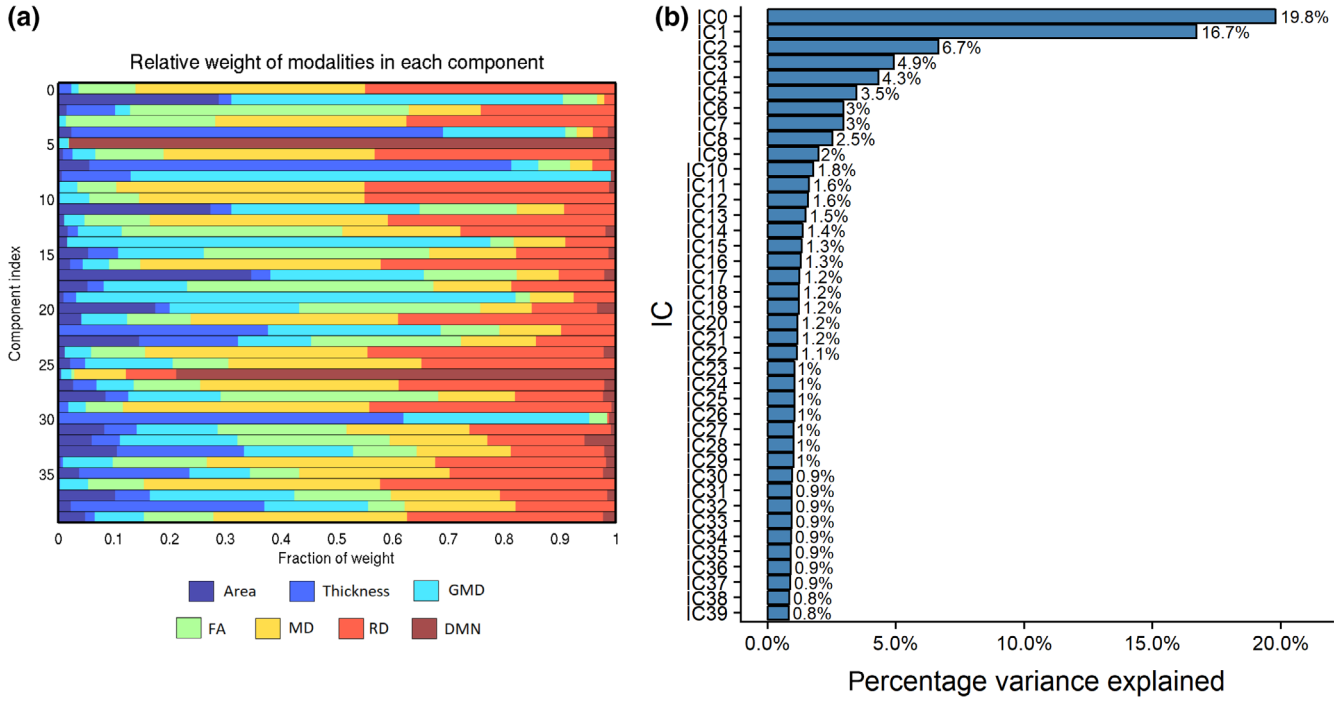

FIGURE 2 (a) The degree of fusing across MRI measures. Specifically, the plot shows the weight (x axis) of each imaging measure (area, thickness, GMD, FA, MD, RD, and DMN) for every LICA component (y axis). (b) The percentage variance explained of each IC, which is used as the sorting criteria of the components in LICA. FA, fractional anisotropy; DMN, default mode network; GMD, gray matter density; LICA, linked independent component analysis; MD, mean diffusivity; MRI, magnetic resonance imaging [Color figure can be viewed at wileyonlinelibrary.com]

Strimmer, 2011). Here, statistical significance was based on RMSE and permutation testing. As a comparison, we also predicted age using the same framework, using residuals of the subject weights after regressing out phase encoding direction in the methods shown above (using Pearson's r instead of Spearman's ρ).

3 | RESULTS

3.1 | Linked independent component analysis

Figure 2a shows the degree of fusion across MRI measures for each component. Figure 2b shows the percentage of the total variance explained by each IC. Most of the components were characterized by region-specific features that were mainly bilateral, with the exception of four global components shown in Figure 3. Briefly, there was no substantial fusion between DMN maps and the other modalities, except for IC26 (Figure S6, Supporting Information).

3.2 | Univariate analyses

Table 2 shows results from linear models testing for main effects of group, age, sex, and symptom load for depression and anxiety on each IC. Table S1 (Supporting Information) shows results from linear models testing for interactions between group or symptom loads for

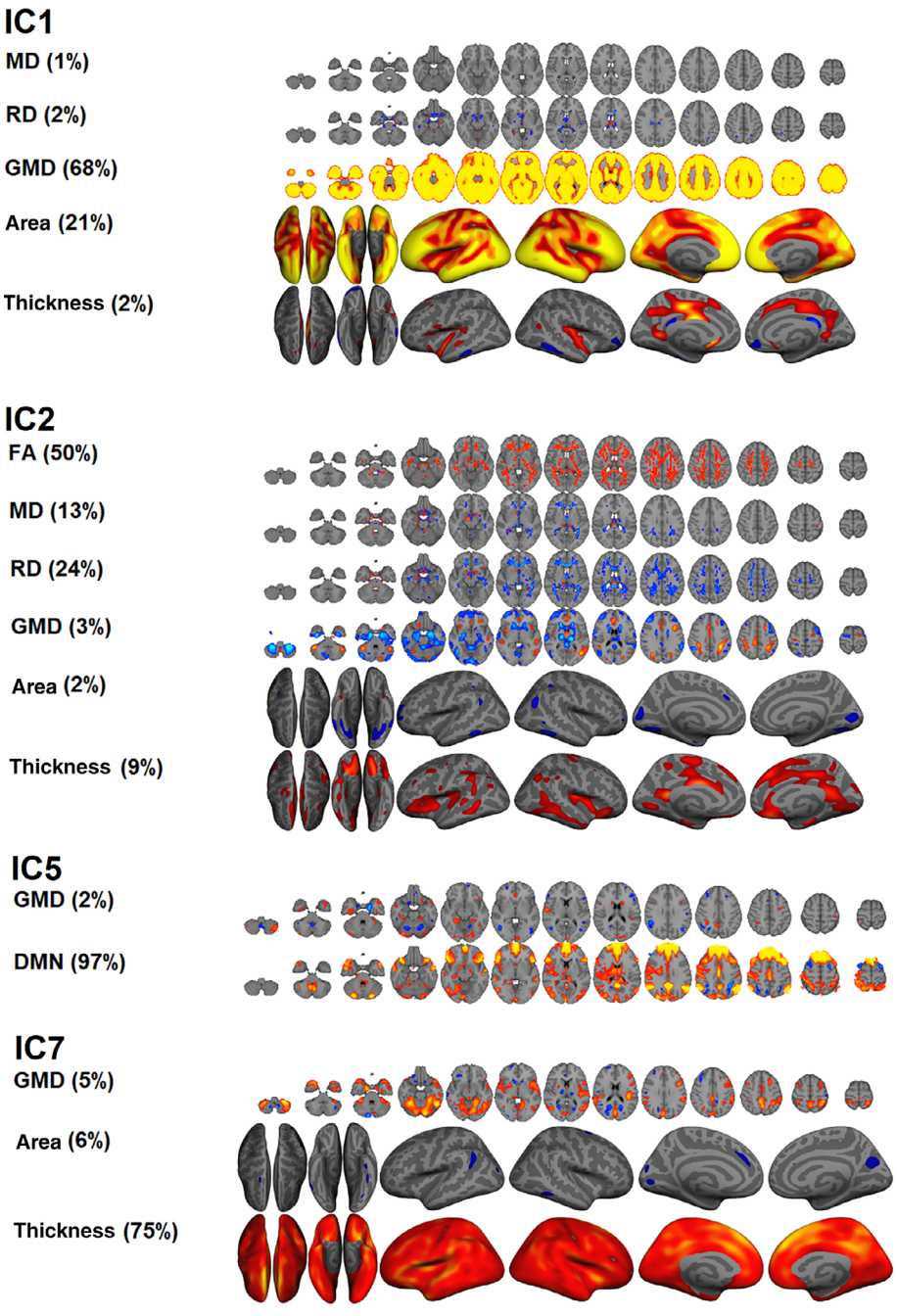

FIGURE 3 ICs that are mainly dominated by global features. For each IC, only measures that have an interpretable spatial pattern are presented. A z-score threshold of > = |3| was used for illustration. For visualization of the skeleton-based ICs, we used tbss_fill. IC1: global gray matter density (GMD) and surface area. IC2: global white matter microstructure. IC5: default mode network (DMN) amplitude. IC7: global thickness [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Main effects of age, sex, group, symptom loads for depression (BDI-II) and anxiety (BAI) on ICs. *p*-Values are FDR corrected

IC	Age (*t*, *p*)	Sex (*t*, *p*)	Group (*t*, *p*)	BDI-II (*t*, *p*)	BAI (*t*, *p*)
IC0	7.69 (<.001)	1.48 (.287)	0.98 (.883)	0.96 (.853)	−0.04 (.973)
IC1	−8.46 (<.001)	10.85 (<.001)	0.32 (.883)	−0.59 (.969)	0.308 (.965)
IC2	−10.21 (<.001)	−6.95 (<.001)	0.57 (.883)	−0.42 (.969)	−0.29 (.965)
IC3	−1.70 (.202)	−1.87 (.208)	0.86 (.883)	1.10 (.853)	1.40 (.965)
IC4	−4.88 (<.001)	1.41 (.307)	0.86 (.883)	−0.14 (.976)	0.33 (.965)
IC5	−6.34 (<.001)	3.45 (.005)	−1.25 (.883)	−0.42 (.969)	−0.75 (.965)
IC6	−3.89 (.001)	0.28 (.841)	1.34 (.883)	−2.23 (.516)	−2.49 (.536)
IC7	−3.56 (.002)	−0.54 (.759)	1.00 (.883)	0.51 (.969)	1.00 (.965)
IC8	−1.64 (.215)	0.35 (.83)	1.32 (.883)	0.31 (.976)	1.06 (.965)
IC9	0.81 (.551)	−1.87 (.208)	−0.62 (.883)	0.55 (.969)	0.35 (.965)
IC10	−0.61 (.62)	−1.76 (.245)	−0.21 (.925)	0.47 (.969)	−0.07 (.973)
IC11	1.34 (.33)	0.40 (.83)	−1.26 (.883)	0.11 (.976)	−0.99 (.965)
IC12	−2.06 (.101)	0.84 (.578)	0.37 (.883)	−1.40 (.853)	−0.94 (.965)
IC13	−5.90 (<.001)	−4.24 (<.001)	−0.04 (.968)	−0.32 (.976)	−0.37 (.965)
IC14	−1.11 (.397)	1.17 (.404)	0.21 (.925)	0.08 (.976)	0.62 (.965)
IC15	0.10 (.932)	5.25 (<.001)	−0.85 (.883)	−0.42 (.969)	−1.08 (.965)
IC16	0.64 (.62)	3.09 (.015)	1.24 (.883)	0.74 (.969)	0.20 (.965)
IC17	−0.97 (.478)	1.27 (.372)	−0.38 (.883)	0.11 (.976)	−0.42 (.965)
IC18	0.51 (.677)	−0.33 (.83)	−0.90 (.883)	−0.540 (.969)	−0.51 (.965)
IC19	−2.68 (.026)	1.25 (.372)	2.84 (.198)	−1.79 (.6)	0.51 (.965)
IC20	−2.47 (.044)	0.15 (.925)	0.72 (.883)	1.02 (.853)	0.11 (.973)
IC21	0.88 (.527)	0.69 (.653)	0.97 (.883)	−0.95 (.853)	−0.42 (.965)
IC22	−0.63 (.62)	−1.71 (.254)	0.97 (.883)	−0.14 (.976)	−0.12 (.973)
IC23	2.21 (.075)	−1.62 (.286)	0.45 (.883)	−1.09 (.853)	−1.29 (.965)
IC24	1.18 (.38)	−0.05 (.96)	0.61 (.883)	0.57 (.969)	0.36 (.965)
IC25	−0.27 (.832)	1.48 (.287)	0.54 (.883)	−0.35 (.976)	0.37 (.965)
IC26	1.20 (.38)	−0.44 (.83)	0.60 (.883)	−1.18 (.853)	−0.40 (.965)
IC27	0.09 (.932)	0.32 (.83)	−0.88 (.883)	−1.45 (.853)	−1.53 (.965)
IC28	0.48 (.68)	1.52 (.287)	0.52 (.883)	1.01 (.853)	0.69 (.965)
IC29	0.80 (.551)	−1.98 (.198)	−0.07 (.968)	0.98 (.853)	0.27 (.965)
IC30	3.64 (.002)	0.72 (.653)	0.52 (.883)	0.63 (.969)	−0.03 (.973)
IC31	−2.99 (.011)	1.52 (.287)	0.36 (.883)	−0.11 (.976)	−0.22 (.965)
IC32	−1.72 (.202)	−2.44 (.078)	0.38 (.883)	1.94 (.542)	1.27 (.965)
IC33	−5.16 (<.001)	−1.00 (.488)	0.33 (.883)	−0.001 (.999)	0.65 (.965)
IC34	−0.72 (.587)	2.30 (.1)	0.10 (.965)	0.57 (.969)	0.20 (.965)
IC35	1.25 (.373)	−0.05 (.96)	0.15 (.952)	1.16 (.853)	1.73 (.965)
IC36	1.44 (.291)	−1.54 (.287)	0.39 (.883)	−2.26 (.516)	−0.74 (.965)
IC37	−1.45 (.291)	−2.70 (.044)	−1.01 (.883)	0.06 (.976)	−0.59 (.965)
IC38	2.25 (.073)	1.13 (.418)	0.52 (.883)	−2.08 (.516)	−1.42 (.965)
IC39	−1.16 (.38)	0.98 (.488)	−1.03 (.883)	−0.25 (.976)	−0.74 (.965)

Abbreviations: BAI, Becks Anxiety Inventory; BDI-II, Becks Depression Inventory; FDR, false discovery rate.

depression or anxiety with age or sex. Briefly, after corrections for multiple comparisons, the analysis revealed no significant associations between ICs and group, nor symptom load for depression and anxiety. There were significant main effects of age and sex on 10 and 5 LICA components (see Figure 4), respectively, but no significant main effects of group, with similar results for the decomposition with 80 components (see Tables S2 and S3, Supporting Information). Figure 3 shows the global LICA components associated with age and sex (see Figure S7, Supporting Information for associations with additional LICA components). Age was negatively associated with IC1,

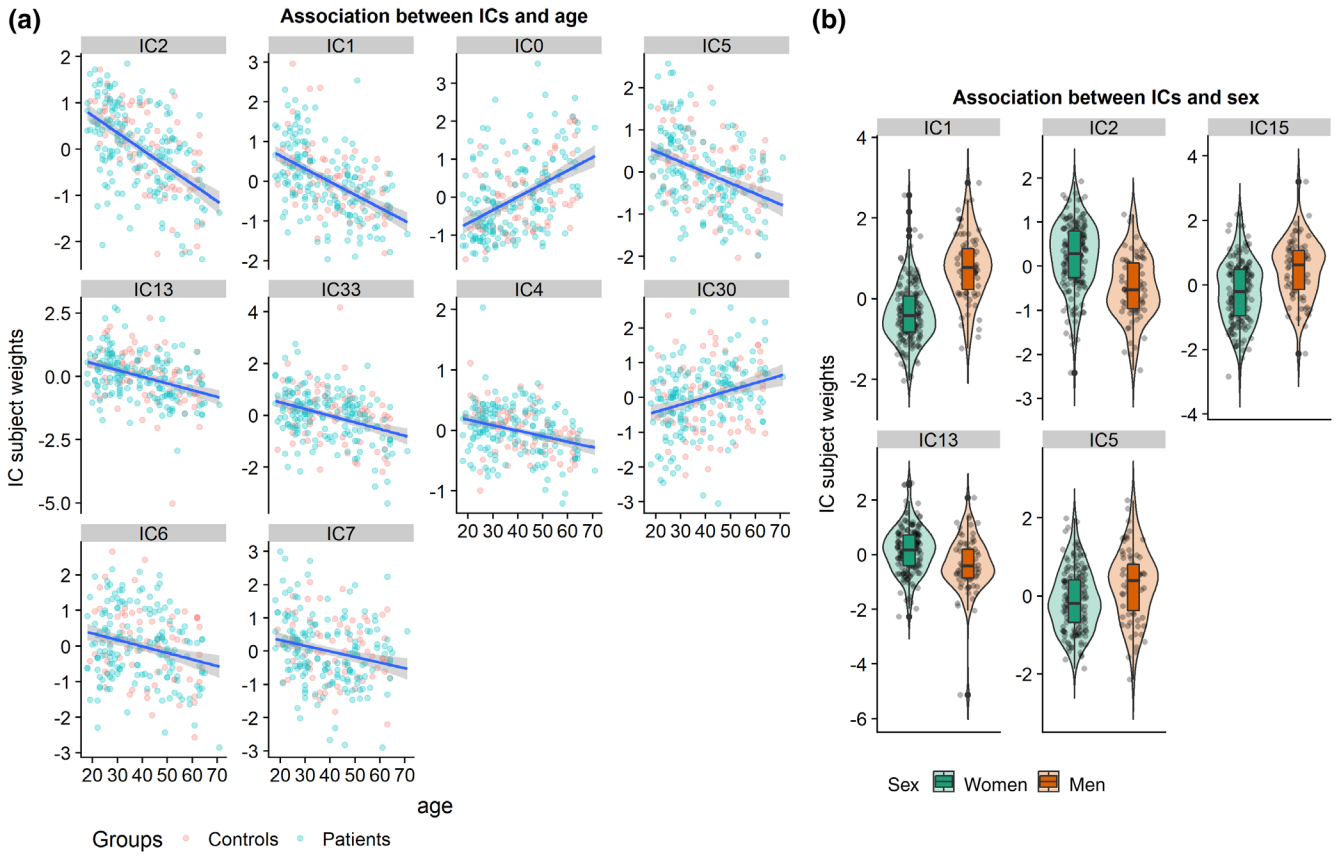

FIGURE 4 Significant effects of age and sex on ICs (false discovery rate [FDR] corrected $p < .01$). (a) Scatter plots of the significant (linear) effects of age on ICs, sorted by the strength of the association. For visualization purposes, we separated based on case–control status. The IC subject weights in the plot have been residualized for group, sex, and phase encoding Direction. B: Violin plots showing the distribution of the subject weights within men and women for each of the components showing a significant main effect of sex. The subject weights in the plot have been residualized for group, age, and phase encoding direction [Color figure can be viewed at wileyonlinelibrary.com]

indicating lower GMD and cortical surface area globally with increasing age, positively associated with IC2, indicating lower FA globally with increasing age, negatively associated with IC5, indicating lower DMN amplitude with increasing age, and negatively associated with IC7, indicating thinner cortex globally with increasing age. The analyses revealed main effects of sex on IC1, indicating larger global surface area and higher GMD in men compared to women, and IC5, indicating that men had higher DMN amplitude. The analyses revealed no significant interaction effects between group or symptom loads for depression or anxiety with age or sex with any of the ICs.

3.3 | Machine learning analyses

Figure 5 shows the results of the machine learning analyses, with the spatial maps of a select few top features for each model shown in Figure S8, Supporting Information. Model performance was low for classifying patients and controls using residualized IC features (AUC = 0.6194, $p = .0614$, accuracy = 0.6169, sensitivity = 0.6991, specificity = 0.4292). The feature importance based on CAT scores identified IC19 as the most important feature for classifying group. IC19 represents a covarying pattern of high GMD in most cerebellar regions, low GMD in cerebellar crus II, and high GMD in the angular gyri.

Model performance was low for predicting depression symptoms using residualized IC features (RMSE = 10.72, $p = .9236$, MAE = 8.498, $R^2 = -.3302$, Spearman's $\rho = 0.009$). IC0 had the highest feature importance based on CAR scores (positive association). IC0 is characterized by a complex covarying pattern including low GMD in temporal regions, the thalamus and cingulate, and low thickness in the cingulate and frontotemporal regions. In terms of white matter diffusion properties, IC0 is characterized by high FA in several pathways including the posterior thalamic radiation and low FA in the anterior thalamic radiation and fornix, with mostly the reverse pattern for MD and RD.

Model performance was low when predicting anxiety symptoms using residualized IC features (RMSE = 8.181, $p = .8946$, MAE = 6.262, $R^2 = -.4240$, Spearman's $\rho = -0.064$). IC6 had the highest feature importance based on CAR scores (negative association). IC6 is mainly characterized by a complex covarying pattern of high FA in the splenium of the corpus callosum, high FA and low MD and RD in the fornix, high MD and RD in the thalamus, in addition to high GMD in the thalamus, and low GMD in hippocampal and amygdala regions. Model performance was slightly lower when regressing out phase

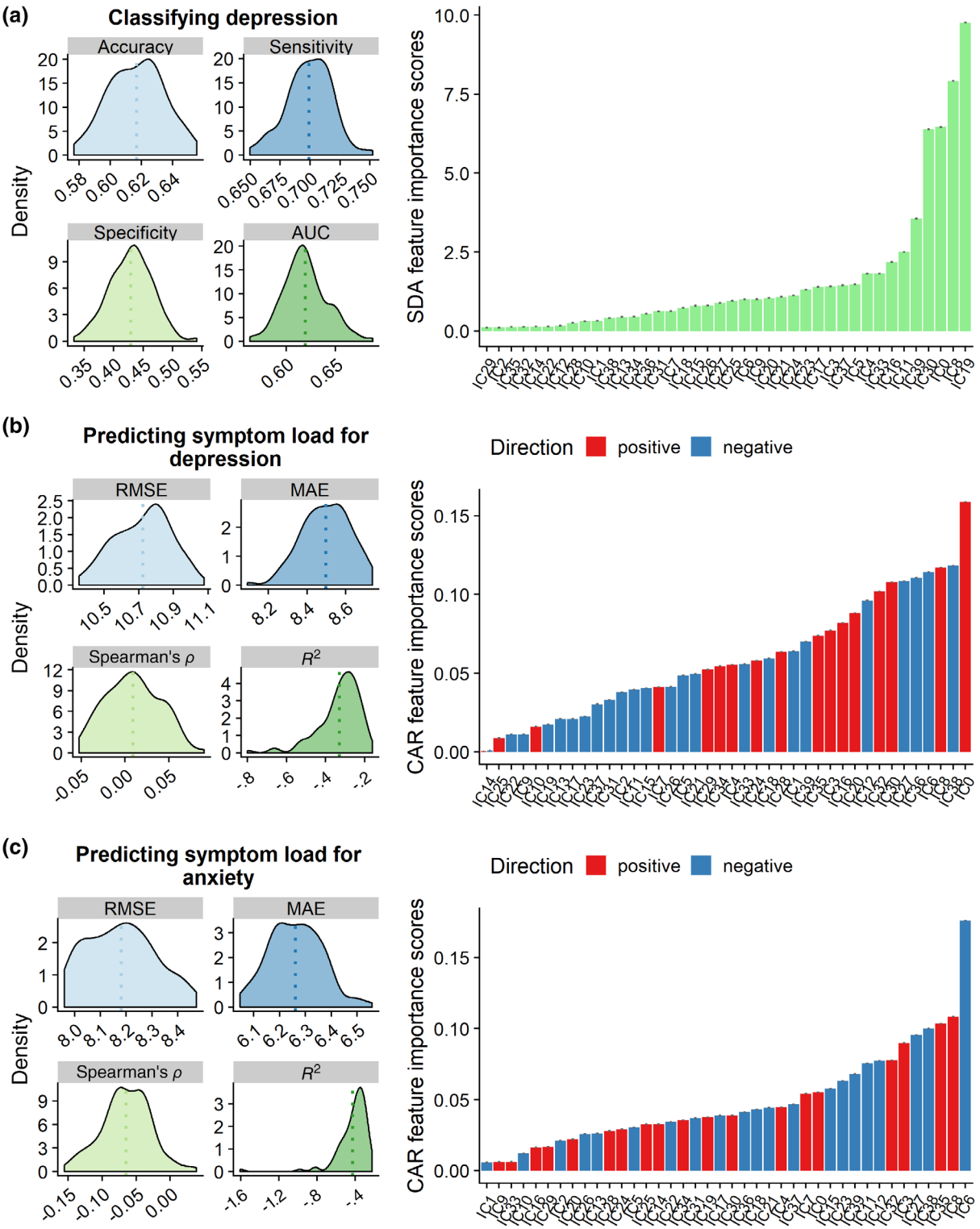

FIGURE 5 The results of the main analyses of the machine learning approach using 10-fold cross-validation across 100 repetitions for (a) classifying group (b) prediction symptom load for depression and (c) symptom load for anxiety. Here, phase encoding direction was only regressed out of the subject weights in IC4, while age and sex were regressed out from the subject weights of all the ICs. The figures on the left show prediction accuracy based on various model performance metrics, with the dotted lines denoting the mean. The barplots on the right show the feature importance of each IC for each model based on CAT scores (a) or CAR-scores (b and c). Direction shows whether the feature is positive or negatively associated with a given model [Color figure can be viewed at wileyonlinelibrary.com]

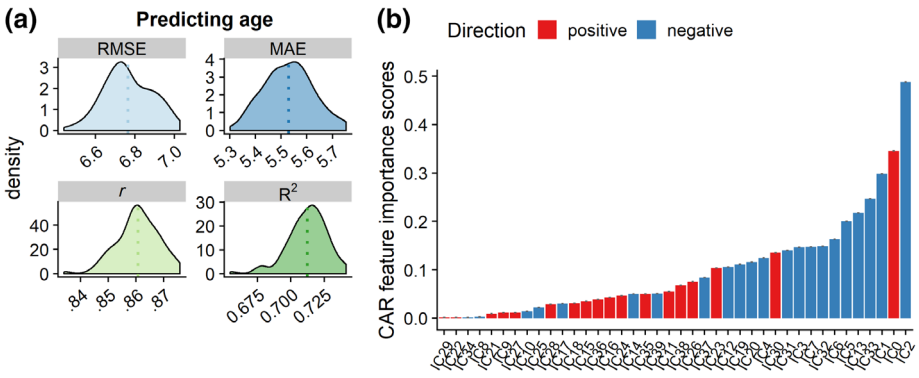

FIGURE 6 Age prediction based on machine learning using 10-fold cross-validation across 100 repetitions. Here, phase encoding direction is regressed out from the subject weights in IC4. (a) Model performance results, with the dotted lines denoting the mean. (b) Feature importance based on CAR scores. Direction shows whether the features is positive or negatively associated with age [Color figure can be viewed at wileyonlinelibrary.com]

encoding from all the IC features, and also suggested a different order of feature importance (see Results and Figure S9, Supporting Information). Using the decomposition with higher model order revealed similar results in terms of feature importance, albeit slightly lower model performance for group classification, and symptom prediction (see Results and Figure S10, Supporting Information). In contrast, model performance was high when predicting age (RMSE = 6.764, $p < .0001$, MAE = 5.530, $R^2 = .7120$, $r = .861$) using residualized features, with feature importance generally in line with the univariate results (see Figure 6). Model performance for predicting age was high but slightly lower when regressing out phase encoding from all the IC features (see Results and Figure S11, Supporting Information), and when using the decomposition with higher model order (see Results and Figure S12, Supporting Information).

Briefly, the results of the supplemental classification analyses accounting for imbalanced group sizes (AUC = 0.632), depression severity (AUC = 0.636), and SSRI medication (AUC = 0.598) were similar to the main analyses. The results of the supplemental analyses accounting for SSRI medication when predicting symptom loads for depression (RMSE = 10.83) and anxiety (RMSE = 8.22) were also similar to the main analyses. Full details of these supplemental analyses can be seen in the Results and Figure S13, Supporting Information.

4 | DISCUSSION

The distributed functional and structural neuroanatomy of complex traits and disorders warrants integrated perspectives and analytical approaches. To this end, we probed the neuronal correlates of depression using multimodal fusion across cortical macrostructure, white matter diffusion properties, and DMN amplitude based on resting-state fMRI. LICA yielded 40 components with various degrees of multimodal involvement and different anatomical distributions, including both global and regionally specific patterns. Univariate analyses revealed strong associations with age and sex for several components' subject weights after multiple comparison correction, but no robust group differences between patients with a history of depression and healthy controls, and no significant interactions between group and sex or between group and age. Likewise, we observed no robust associations with symptom loads for depression or anxiety, nor interactions with age or sex. In line with the univariate analyses, the machine learning approach revealed overall low prediction accuracy for group classification and prediction of symptom loads for depression and anxiety, but high prediction accuracy for age.

Our univariate analyses revealed no main effects of history of depression or symptom load for depression on any of the LICA components. The machine learning analyses here revealed overall low predictive value both for case–control status and symptoms of depression and anxiety, which is generally in line with the univariate analyses and an increasing body of literature suggesting small differences in brain structure between patients with MDD and healthy controls (Schmaal et al., 2016; Schmaal et al., 2017; Varoquaux, 2018; Wolfers, Buitelaar, Beckmann, Franke, & Marquand, 2015). While considering the overall low performance, the most important feature for classifying patients with a history of depression from healthy controls was a component encompassing covarying patterns of both high and low GMD in cerebellar regions (IC19). There is some evidence that depression is linked to cerebellar regions that communicate with networks related to cognitive and introspective processing (Depping, Schmitgen, Kubera, & Wolf, 2018). Cerebellar structural characteristics have recently been demonstrated to rank among the most sensitive brain features when comparing adult patients with schizophrenia and healthy controls (Moberget et al., 2018), and also for predicting psychiatric symptoms in youths (Moberget et al., 2019). The most important feature for predicting depression symptoms was IC0, which involves complex covarying patterns of low GMD and cortical thickness in mainly temporal but also frontal regions. This pattern is largely in line with previous research (Lai, 2013; Schmaal et al., 2017; Suh et al., 2019). IC0 also encompassed high FA and low MD and RD in interhemispheric connections and frontal–striatal thalamic pathways, in line with previous studies, albeit in the opposite direction (Chen et al., 2016). One study reported a positive association between symptom load for depression and FA in the thalamus (Osoba et al., 2013), and another study suggested this association (although in the opposite direction) is related to late onset MDD, especially in the corpus callosum (Cheng et al., 2014). However, while the implicated brain patterns are in line with previous reports, the overall poor model performance for classifying cases from controls, and predicting symptom loads for depression and anxiety warrants caution when interpreting feature importance.

Higher age was related to lower global cortical thickness (IC7), in line with previous studies (e.g., Fjell et al., 2015). As hypothesized, higher age was also associated with lower global volume and smaller surface area (IC1), similar to previous studies using LICA (Doan, Engvig, Zaske, et al., 2017; Douaud et al., 2014), and a consistent finding in lifespan studies. Additionally, advancing age was negatively associated with IC2, indicating decreased FA globally, but also increased RD and to some extent MD, consistent with the aging literature (Davis et al., 2009; Sexton et al., 2014; Westlye et al., 2010). Higher age was associated with IC5, reflecting age-related decreases in DMN amplitude, in line with previous research (Damoiseaux et al., 2008; Mevel et al., 2013; Mowinckel, Espeseth, & Westlye, 2012; Razlighi et al., 2014; Vidal-Piñeiro et al., 2014). We observed high prediction accuracy for age in the machine learning approach. This finding supports the potential utility of LICA in estimating the gap between chronological and biological age (i.e., brain age gap).

Men had larger global brain volume and surface area than women (IC1), which is consistent with previous studies (e.g., Ritchie et al., 2018). Additionally, women had higher subject weights in IC13, reflecting lower FA in the corticospinal tract, portions of the superior longitudinal fasciculi and posterior thalamic radiation compared to men, generally in line with a large-scale UK Biobank study (Ritchie et al., 2018). We also found that men had greater DMN amplitude (IC5) than women, which adds to previous inconclusive findings (Mowinckel et al., 2012; Weissman-Fogel, Moayedi, Taylor, Pope, & Davis, 2010) and contrasts a previous report suggesting effects in the opposite direction (Jamadar et al., 2018).

In general, our analyses did not provide support for our hypothesis that a history of depression and symptoms of depression interact with age-related trajectories or sex differences of the LICA components. To the best of our knowledge, although studies have found specific cortical abnormalities related to adults with MDD, adolescents with MDD (Schmaal et al., 2017) and age at onset of depression (Ho et al., 2019), few or no studies have reported age-by-group interactions. A recent large-scale analysis of regional brain age estimations based on brain morphometry reported increased temporal lobe brain age in patients with MDD compared to matched healthy peers (Kaufmann et al., in press), with a relatively small effect size. Although there have been some early reports of a sex-by-group interaction in hippocampal volumes (e.g., Frodl et al., 2002), the recent large-scale ENIGMA MDD study reported no sex-by-group interactions in any subcortical volumes (Schmaal et al., 2016). Another recent study identified sex-by-group interactions using VBM, including higher GMD in the left cerebellum of male patients only, and lower GMD in the dorsal medial prefrontal gyrus in female patients only (Yang et al., 2017). However, the sample size was relatively small (less than 100 in the patient and control group each) which may affect the reproducibility of these findings. Despite separate reports of a link between resting-state DMN connectivity and rumination in depression (e.g., Hamilton, Farmer, Fogelman, & Gotlib, 2015) and evidence of sex-differences in rumination among adolescents (e.g., Jose & Brown, 2008), no studies have reported a sex-by-group interaction on DMN functional connectivity or amplitude.

The lack of positive findings in the univariate analyses and low predictive accuracy in the machine learning approach can be attributed to at least two factors. First, as illustrated by the large-scale ENIGMA studies (Ho et al., 2019; Schmaal et al., 2016; Schmaal et al., 2017), the effect sizes in neuroimaging studies of mental disorders and depression are overall small (Paulus & Thompson, 2019). Similarly, small sample sizes may contribute to inflated estimates of prediction accuracy in machine learning approaches (Wolfers et al., 2015). This is one possible explanation why other multimodal fusion studies of depression have reached higher prediction accuracies for classifying patients with depression from healthy controls (He et al., 2017; Ramezani et al., 2014; Yang et al., 2018), with patient groups consisting of no more than 60 individuals. Secondly, and also related to the small effect sizes, mental disorders including depression are clinically highly heterogeneous. As an example, Müller et al. (2017) partially attribute the lack of convergence in their meta-analysis of activation-based fMRI experiments involving 1,058 MDD patients to clinical heterogeneity. As a result, future research probing the neurobiology of depression should aim for large sample sizes (Rutledge, Chekroud, & Huys, 2019), and more importantly, stratifying patients (Feczko et al., 2019) with depression at the individual level. Pursuant to this, there has been considerable interest in identifying clinically relevant subgroups based on brain imaging, with initially encouraging results (Drysdale et al., 2017). However, the robustness and generalizability of such studies have been brought into question (Dinga et al., 2019), which may be partly due to substantial brain heterogeneity within groups, which has been illustrated in terms of morphometry in schizophrenia (Alnæs et al., 2019). Alternatively, dimensional measures such as brain age prediction (Kaufmann et al., In press) and normative modeling (Marquand et al., 2019; Marquand, Rezek, Buitelaar, & Beckmann, 2016) have shown promising results in elucidating brain heterogeneity in mental disorders such as schizophrenia (Wolfers et al., 2018) and attention deficit/hyperactivity disorder (Wolfers et al., 2019). An important consideration is the multicausal nature of depression, which means that it is likely that including other measures such as psychosocial, cognitive, and genetic factors would increase the explained variance, beyond brain imaging by itself.

The current findings should be considered in light of relevant limitations associated with statistical power and study design. The low number of severely depressed patients relative to remitted patients may have influenced the sensitivity and specificity of the machine learning approach, in particular for predicting symptom loads of depression and anxiety. However, supplemental analyses excluding patients with minimal depression yielded similar results, supporting the notion of depression as a heterogeneous mental disorder. The varied current use of antidepressants in the patient group may have impacted the classification accuracy, although it is undoubtedly difficult to get large samples of nonmedicated patients. However, supplemental analyses excluding patients using medications did not alter their results in the aforementioned meta-analysis of activation-based fMRI experiments in depression (Müller et al., 2017). Additionally, the confounding of SSRI use and depression severity in the classification of cases and controls were addressed in supplemental analyses and revealed that results remained the same.

One weakness of this study is that the change in phase encoding direction may have introduced systematic differences in the MRI signal. However, only one component (IC4 in both decompositions) was strongly sensitive to phase encoding direction, suggesting that the remaining components were largely unaffected. Furthermore, we accounted for phase encoding direction in both the univariate and the machine learning analyses. This along with previous studies (Doan, Engvig, Persson, et al., 2017; Doan, Engvig, Zaske, et al., 2017; Groves et al., 2012) provide additional evidence that LICA is a promising tool to account for various scanner effects, particularly relevant for multisite and longitudinal studies. We used a model order of 40 for LICA decomposition in the main analyses. Even though we did find similar feature importance ranking in the decomposition with higher model order, prediction accuracy was slightly lower across all models. Although there is no consensus on the optimal model order, this may imply that we were modeling more noise components in the higher model order decomposition, which was also supported by the number of discarded components due to dominance by a single subject.

In conclusion, based on fusion of structural, diffusion-weighted, and resting-state fMRI data from 241 individuals with or without a history of depression, we identified multimodal and modality specific components that revealed strong associations with age and sex. None of the components showed significant association with categorical or dimensional measures of depression, nor any interaction effects with age and sex. Similarly, machine learning revealed low prediction accuracy for classifying patients from controls and predicting symptom loads. This study supports accumulating evidence of small effect sizes when comparing brain imaging features between patients with a history of depression and healthy controls, and indicates the need for more precise methods of stratifying individuals with depression, as well as large samples sizes.

ACKNOWLEDGMENTS

This work was supported by the South-Eastern Norway Regional Health Authority (2014097, 2015073, 2015052), the Research Council of Norway (249795, 229135, 175387/V50), the European Research Council (ERC StG 802998), and the Department of Psychology, University of Oslo. The permutation testing was performed using resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway.

CONFLICT OF INTERESTS

N.I.L. has previously received consultancy fees and travel expenses from Lundbeck. All the other authors have nothing to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Luigi A. Maglanoc https://orcid.org/0000-0003-2556-0779
Lars T. Westlye https://orcid.org/0000-0001-8644-956X

REFERENCES

Ahdesmäki, M., & Strimmer, K. (2010). Feature selection in omics prediction problems using cat scores and false nondiscovery rate control. *The Annals of Applied Statistics*, *4*(1), 503–519. https://doi.org/10.1214/09-AOAS277

Alnæs, D., Kaufmann, T., Doan, N. T., Córdova-Palomera, A., Wang, Y., Bettella, F., … Westlye, L. T. (2018). Association of heritable cognitive ability and psychopathology with white matter properties in children and adolescents. *JAMA Psychiatry*, *75*(3), 287–295. https://doi.org/10.1001/jamapsychiatry.2017.4277

Alnæs, D., Kaufmann, T., van der Meer, D., Córdova-Palomera, A., Rokicki, J., Moberget, T., … Westlye, L. T. (2019). Brain heterogeneity in schizophrenia and its association with polygenic risk. *JAMA Psychiatry*, *76*, 739–748. https://doi.org/10.1001/jamapsychiatry.2019.0257

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Beck, A. (1996). *Manual for the Beck depression inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, A., & Steer, R. (1993). *Beck anxiety inventory manual*. San Antonio, TX: Psychological Corporation.

Chen, G., Hu, X., Li, L., Huang, X., Lui, S., Kuang, W., … Gong, Q. (2016). Disorganization of white matter architecture in major depressive disorder: A meta-analysis of diffusion tensor imaging with tract-based spatial statistics. *Scientific Reports*, *6*, 21825. https://doi.org/10.1038/srep21825

Cheng, Y., Xu, J., Yu, H., Nie, B., Li, N., Luo, C., … Xu, X. (2014). Delineation of early and later adult onset depression by diffusion tensor imaging. *PLoS One*, *9*(11), e112307. https://doi.org/10.1371/journal.pone.0112307

Cuthbert, B. N., & Insel, T. R. (2012). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *American Journal of Psychiatry*, *11*(1), 126. https://doi.org/10.1186/1741-7015-11-126

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Damoiseaux, J. S., Beckmann, C. F., Arigita, E. J. S., Barkhof, F., Scheltens, P., Stam, C. J., … Rombouts, S. A. R. B. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(8), 1856–1864. https://doi.org/10.1093/cercor/bhm207

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., & Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *NeuroImage*, *46*(2), 530–541. https://doi.org/10.1016/j.neuroimage.2009.01.068

Depping, M. S., Schmitgen, M. M., Kubera, K. M., & Wolf, R. C. (2018). Cerebellar contributions to major depression. *Frontiers in Psychiatry*, *9*, 634. https://doi.org/10.3389/fpsyt.2018.00634

Dinga, R., Schmaal, L., Penninx, B. W. J. H., van Tol, M. J., Veltman, D. J., van Velzen, L., … Marquand, A. F. (2019). Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). *NeuroImage: Clinical*, *22*, 101796. https://doi.org/10.1016/j.nicl.2019.101796

Doan, N. T., Engvig, A., Persson, K., Alnæs, D., Kaufmann, T., Rokicki, J., … Westlye, L. T. (2017). Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer's disease spectrum. *Scientific Reports*, *7*, 45131. https://doi.org/10.1038/srep45131

Doan, N. T., Engvig, A., Zaske, K., Persson, K., Lund, M. J., Kaufmann, T., … Alzheimer's Disease Neuroimaging Initiative. (2017). Distinguishing early and late brain aging from the Alzheimer's disease spectrum: Consistent morphological patterns across independent samples. *NeuroImage*, *158*, 282–295. https://doi.org/10.1016/j.neuroimage.2017.06.070

Douaud, G., Groves, A. R., Tamnes, C. K., Westlye, L. T., Duff, E. P., Engvig, A., … Johansen-Berg, H. (2014). A common brain network links development, aging, and vulnerability to disease. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(49), 17648–17653. https://doi.org/10.1073/pnas.1410378111

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28–38. https://doi.org/10.1038/nm.4246

Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A. M., Nigg, J. T., & Fair, D. A. (2019). The heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends in Cognitive Sciences*, *23*, 584–601. https://doi.org/10.1016/j.tics.2019.03.009

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396

Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., … Walhovd, K. B. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(50), 15462–15467. https://doi.org/10.1073/pnas.1508831112

Francx, W., Llera, A., Mennes, M., Zwiers, M. P., Faraone, S. V., Oosterlaan, J., … Beckmann, C. F. (2016). Integrated analysis of gray and white matter alterations in attention-deficit/hyperactivity disorder. *NeuroImage. Clinical*, *11*, 357–367. https://doi.org/10.1016/j.nicl.2016.03.005

Fried, E. I., & Kievit, R. A. (2016). The volumes of subcortical regions in depressed and healthy individuals are strikingly similar: A reinterpretation of the results by Schmaal et al. *Molecular Psychiatry*, *21*(6), 724–725. https://doi.org/10.1038/mp.2015.199

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, *172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Friedrich, M. J. (2017). Depression is the leading cause of disability around the world. *JAMA*, *317*(15), 1517. https://doi.org/10.1001/jama.2017.3826

Frodl, T., Meisenzahl, E. M., Zetzsche, T., Born, C., Groll, C., Jäger, M., … Möller, H.-J. (2002). Hippocampal changes in patients with a first episode of major depression. *The American Journal of Psychiatry*, *159*(7), 1112–1118. https://doi.org/10.1176/appi.ajp.159.7.1112

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., … Smith, S. M. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*, *95*, 232–247. https://doi.org/10.1016/j.neuroimage.2014.03.034

Groves, A. R., Beckmann, C. F., Smith, S. M., & Woolrich, M. W. (2011). Linked independent component analysis for multimodal data fusion. *NeuroImage*, *54*(3), 2198–2217. https://doi.org/10.1016/j.neuroimage.2010.09.073

Groves, A. R., Smith, S. M., Fjell, A. M., Tamnes, C. K., Walhovd, K. B., Douaud, G., … Westlye, L. T. (2012). Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. *NeuroImage*, *63*(1), 365–380. https://doi.org/10.1016/j.neuroimage.2012.06.038

Hamilton, J. P., Farmer, M., Fogelman, P., & Gotlib, I. H. (2015). Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience. *Biological Psychiatry*, *78*(4), 224–230. https://doi.org/10.1016/j.biopsych.2015.02.020

He, H., Sui, J., Du, Y., Yu, Q., Lin, D., Yang, J., … Calhoun, V. D. (2017). Co-altered functional networks and brain structure in unmedicated patients with bipolar and major depressive disorders. *Brain Structure & Function*, *222*(9), 4051–4064. https://doi.org/10.1007/s00429-017-1451-x

Ho, T. C., Gutman, B., Pozzi, E., Grabe, H. J., Hosten, N., Wittfeld, K., … Schmaal, L. (2019). Subcortical shape alterations in major depressive disorder: Findings from the ENIGMA major depressive disorder working group. *BioRxiv*. https://doi.org/10.1101/534370

Husain, M. M., Rush, A. J., Sackeim, H. A., Wisniewski, S. R., McClintock, S. M., Craven, N., … Hauger, R. (2005). Age-related characteristics of depression: A preliminary STAR*D report. *The American Journal of Geriatric Psychiatry*, *13*(10), 852–860. https://doi.org/10.1176/appi.ajgp.13.10.852

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, *10*(3), 626–634. https://doi.org/10.1109/72.761722

Insel, T. R. (2014). The NIMH research domain criteria (RDoC) project: Precision medicine for psychiatry. *American Journal of Psychiatry*, *171*(4), 395–397. https://doi.org/10.1176/appi.ajp.2014.14020138

Insel, T. R. (2015). Brain disorders? Precisely. *Science*, *348*(6234), 499–500. https://doi.org/10.1126/science.aab2358

Jamadar, S. D., Sforazzini, F., Raniga, P., Ferris, N. J., Paton, B., Bailey, M. J., … ASPREE Investigator Group. (2018). Sexual dimorphism of resting-state network connectivity in healthy ageing. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*. https://doi.org/10.1093/geronb/gby004. [Epub ahead of print]

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841.

Johansson, R., Carlbring, P., Heedman, Å., Paxling, B., & Andersson, G. (2013). Depression, anxiety and their comorbidity in the Swedish general population: Point prevalence and the effect on health-related quality of life. *PeerJ*, *1*, e98. https://doi.org/10.7717/peerj.98

Jose, P. E., & Brown, I. (2008). When does the gender difference in rumination begin? Gender and age differences in the use of rumination by adolescents. *Journal of Youth and Adolescence*, *37*(2), 180–192. https://doi.org/10.1007/s10964-006-9166-y

Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, *24*(1), 1–10. https://doi.org/10.1016/0165-1684(91)90079-X

Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: Meta-analysis of resting-state functional connectivity. *JAMA Psychiatry*, *72*(6), 603–611. https://doi.org/10.1001/jamapsychiatry.2015.0071

Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., … Westlye, L. T. (in press). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*.

Lai, C.-H. (2013). Gray matter volume in major depressive disorder: A meta-analysis of voxel-based morphometry studies. *Psychiatry Research: Neuroimaging*, *211*(1), 37–46. https://doi.org/10.1016/j.pscychresns.2012.06.006

Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J. L. M., … Penninx, B. W. J. H. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: The Netherlands study of depression and anxiety (NESDA). *The Journal of Clinical Psychiatry*, *72*(3), 341–348. https://doi.org/10.4088/JCP.10m06176blu

Maglanoc, L. A., Landrø, N. I., Jonassen, R., Kaufmann, T., Córdova-Palomera, A., Hilland, E., & Westlye, L. T. (2019). Data-driven clustering reveals a link between symptoms and functional brain connectivity in depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(1), 16–26. https://doi.org/10.1016/j.bpsc.2018.05.005

Marcus, S. M., Young, E. A., Kerber, K. B., Kornstein, S., Farabaugh, A. H., Mitchell, J., … Rush, A. J. (2005). Gender differences in depression: Findings from the STAR*D study. *Journal of Affective Disorders*, *87* (2–3), 141–150. https://doi.org/10.1016/j.jad.2004.09.008

Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., & Beckmann, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry*, *1*. https://doi.org/10.1038/s41380-019-0441-1. [Epub ahead of print]

Marquand, A. F., Rezek, I., Buitelaar, J., & Beckmann, C. F. (2016). Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. *Biological Psychiatry*, *80*(7), 552–561. https://doi.org/10.1016/j.biopsych.2015.12.023

Mevel, K., Landeau, B., Fouquet, M., La Joie, R., Villain, N., Mézenge, F., … Chételat, G. (2013). Age effect on the default mode network, inner thoughts, and cognitive abilities. *Neurobiology of Aging*, *34*(4), 1292–1301. https://doi.org/10.1016/j.neurobiolaging.2012.08.018

Moberget, T., Alnæs, D., Kaufmann, T., Doan, N. T., Córdova-Palomera, A., Norbom, L. B., … Westlye, L. T. (2019). Cerebellar gray matter volume is associated with cognitive function and psychopathology in adolescence. *Biological Psychiatry*, *86*, 65–75. https://doi.org/10.1016/j.biopsych.2019.01.019

Moberget, T., Doan, N. T., Alnæs, D., Kaufmann, T., Córdova-Palomera, A., Lagerberg, T. V., … Westlye, L. T. (2018). Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: A multisite mega-analysis of 983 patients and 1349 healthy controls. *Molecular Psychiatry*, *23*(6), 1512–1520. https://doi.org/10.1038/mp.2017.106

Mowinckel, A. M., Espeseth, T., & Westlye, L. T. (2012). Network-specific effects of age and in-scanner subject motion: A resting-state fMRI study of 238 healthy adults. *NeuroImage*, *63*(3), 1364–1373. https://doi.org/10.1016/j.neuroimage.2012.08.004

Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F., & Tendolkar, I. (2015). Resting-state functional connectivity in major depressive disorder: A review. *Neuroscience and Biobehavioral Reviews*, *56*, 330–344. https://doi.org/10.1016/j.neubiorev.2015.07.014

Müller, V. I., Cieslik, E. C., Serbanescu, I., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2017). Altered brain activity in unipolar depression revisited: Meta-analyses of neuroimaging studies. *JAMA Psychiatry*, *74* (1), 47–55. https://doi.org/10.1001/jamapsychiatry.2016.2783

Nickerson, L. D., Smith, S. M., Öngür, D., & Beckmann, C. F. (2017). Using dual regression to investigate network shape and amplitude in functional connectivity analyses. *Frontiers in Neuroscience*, *11*, 115. https://doi.org/10.3389/fnins.2017.00115. eCollection 2017.

Osoba, A., Hänggi, J., Li, M., Horn, D. I., Metzger, C., Eckert, U., … Walter, M. (2013). Disease severity is correlated to tract specific changes of fractional anisotropy in MD and CM thalamus—A DTI study in major depressive disorder. *Journal of Affective Disorders*, *149* (1–3), 116–128. https://doi.org/10.1016/j.jad.2012.12.026

Paulus, M. P., & Thompson, W. K. (2019). The challenges and opportunities of small effects: The new normal in academic psychiatry. *JAMA Psychiatry*, *76*(4), 353–354. https://doi.org/10.1001/jamapsychiatry.2018.4540

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ramezani, M., Abolmaesumi, P., Tahmasebi, A., Bosma, R., Tong, R., Hollenstein, T., … Johnsrude, I. (2014). Fusion analysis of first episode depression: Where brain shape deformations meet local composition of tissue. *NeuroImage: Clinical*, *7*, 114–121. https://doi.org/10.1016/j.nicl.2014.11.016

Razlighi, Q. R., Habeck, C., Steffener, J., Gazes, Y., Zahodne, L. B., MacKay-Brandt, A., & Stern, Y. (2014). Unilateral disruptions in the default network with aging in native space. *Brain and Behavior*, *4*(2), 143–157. https://doi.org/10.1002/brb3.202

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., … Deary, I. J. (2018). Sex differences in the adult human brain: Evidence from 5216 UKbiobank participants. *Cerebral Cortex*, *28*(8), 2959–2975. https://doi.org/10.1093/cercor/bhy109

Roalf, D. R., Quarmley, M., Elliott, M. A., Satterthwaite, T. D., Vandekar, S. N., Ruparel, K., … Gur, R. E. (2016). The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *NeuroImage*, *125*, 903–919. https://doi.org/10.1016/j.neuroimage.2015.10.068

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. https://doi.org/10.1186/1471-2105-12-77

Rutledge, R. B., Chekroud, A. M., & Huys, Q. J. (2019). Machine learning and big data in psychiatry: Toward clinical applications. *Current Opinion in Neurobiology*, *55*, 152–159. https://doi.org/10.1016/j.conb.2019.02.006

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*, *90*, 449–468. https://doi.org/10.1016/j.neuroimage.2013.11.046

Schäfer, J., & Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, *4(1)*. https://doi.org/10.2202/1544-6115.1175

Schmaal, L., Hibar, D. P., Sämann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., … Veltman, D. J. (2017). 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. *Molecular Psychiatry*, *22*(6), 900–909. https://doi.org/10.1038/mp.2016.60

Schmaal, L., Veltman, D. J., van Erp, T. G. M., Sämann, P. G., Frodl, T., Jahanshad, N., … Hibar, D. P. (2016). Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA major depressive disorder working group. *Molecular Psychiatry*, *21*(6), 806–812. https://doi.org/10.1038/mp.2015.69

Sexton, C. E., Walhovd, K. B., Storsve, A. B., Tamnes, C. K., Westlye, L. T., Johansen-Berg, H., & Fjell, A. M. (2014). Accelerated changes in white matter microstructure during aging: A longitudinal diffusion tensor imaging study. *The Journal of Neuroscience*, *34*(46), 15425–15436. https://doi.org/10.1523/JNEUROSCI.0203-14.2014

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*(Suppl. 20), 22–33 quiz 34–57.

Suh, J. S., Schneider, M. A., Minuzzi, L., MacQueen, G. M., Strother, S. C., Kennedy, S. H., & Frey, B. N. (2019). Cortical thickness in major depressive disorder: A systematic review and meta-analysis. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *88*, 287–302. https://doi.org/10.1016/j.pnpbp.2018.08.008

Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*, *180*(Pt. A), 68–77. https://doi.org/10.1016/j.neuroimage.2017.06.061

Vidal-Piñeiro, D., Valls-Pedret, C., Fernández-Cabello, S., Arenaza-Urquijo, E. M., Sala-Llonch, R., Solana, E., … Bartrés-Faz, D. (2014). Decreased default mode network connectivity correlates with age-associated structural and cognitive changes. *Frontiers in Aging Neuroscience*, *6*, 256. https://doi.org/10.3389/fnagi.2014.00256

Weissman-Fogel, I., Moayedi, M., Taylor, K. S., Pope, G., & Davis, K. D. (2010). Cognitive and default-mode resting state networks: Do male and female brains "rest" differently? *Human Brain Mapping*, *31*(11), 1713–1726. https://doi.org/10.1002/hbm.20968

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., … Fjell, A. M. (2010). Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(9), 2055–2068. https://doi.org/10.1093/cercor/bhp280

Wolfers, T., Arenas, A. L., Onnink, A. M. H., Dammers, J., Hoogman, M., Zwiers, M. P., … Beckmann, C. F. (2017). Refinement by integration: Aggregated effects of multimodal imaging markers on adult ADHD. *Journal of Psychiatry & Neuroscience*, *42*(6), 386–394. https://doi.org/10.1503/jpn.160240

Wolfers, T., Beckmann, C. F., Hoogman, M., Buitelaar, J. K., Franke, B., & Marquand, A. F. (2019). Individual differences v. the average patient: Mapping the heterogeneity in ADHD using normative models. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291719000084. [Epub ahead of print]

Wolfers, T., Buitelaar, J. K., Beckmann, C. F., Franke, B., & Marquand, A. F. (2015). From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neuroscience and Biobehavioral Reviews*, *57*, 328–349. https://doi.org/10.1016/j.neubiorev.2015.08.001

Wolfers, T., Doan, N. T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., … Marquand, A. F. (2018). Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry*, *75*(11), 1146–1155. https://doi.org/10.1001/jamapsychiatry.2018.2467

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Retrieved from www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/

Wu, Z.-M., Llera, A., Hoogman, M., Cao, Q.-J., Zwiers, M. P., Bralten, J., … Wang, Y.-F. (2019). Linked anatomical and functional brain alterations in children with attention-deficit/hyperactivity disorder. *NeuroImage: Clinical*, *23*, 101851. https://doi.org/10.1016/j.nicl.2019.101851

Yan, C.-G., Chen, X., Li, L., Castellanos, F. X., Bai, T.-J., Bo, Q.-J., … Zang, Y.-F. (2019). Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, *201900390*, 9078–9083. https://doi.org/10.1073/pnas.1900390116

Yang, J., Yin, Y., Zhang, Z., Long, J., Dong, J., Zhang, Y., … Yuan, Y. (2018). Predictive brain networks for major depression in a semi-multimodal fusion hierarchical feature reduction framework. *Neuroscience Letters*, *665*, 163–169. https://doi.org/10.1016/j.neulet.2017.12.009

Yang, X., Peng, Z., Ma, X., Meng, Y., Li, M., Zhang, J., … Ma, X. (2017). Sex differences in the clinical characteristics and brain gray matter volume alterations in unmedicated patients with major depressive disorder. *Scientific Reports*, *7*, 2515. https://doi.org/10.1038/s41598-017-02828-4

Zuber, V., & Strimmer, K. (2011). High-dimensional regression and variable selection using CAR scores. *Statistical Applications in Genetics and Molecular Biology*, *10*(1). https://doi.org/10.2202/1544-6115.1730

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maglanoc LA, Kaufmann T, Jonassen R, et al. Multimodal fusion of structural and functional brain imaging in depression using linked independent component analysis. *Hum Brain Mapp*. 2019; 1–15. https://doi.org/10.1002/hbm.24802