

RESEARCH

Open Access



# AI-based dimensional neuroimaging system for characterizing heterogeneity in brain structure and function in major depressive disorder: COORDINATE-MDD consortium design and rationale

Cynthia H. Y. Fu<sup>1,2\*</sup>, Guray Erus<sup>3</sup>, Yong Fan<sup>3</sup>, Mathilde Antoniadou<sup>3</sup>, Danilo Arnone<sup>2,4</sup>, Stephen R. Arnott<sup>5</sup>, Taolin Chen<sup>6,7</sup>, Ki Sueng Choi<sup>8</sup>, Cherise Chin Fatt<sup>9</sup>, Benicio N. Frey<sup>10,11</sup>, Vibe G. Frokjaer<sup>12,13,14</sup>, Melanie Ganz<sup>12,15</sup>, Jose Garcia<sup>3</sup>, Beata R. Godlewska<sup>16,17</sup>, Stefanie Hassel<sup>18,19</sup>, Keith Ho<sup>20</sup>, Andrew M. McIntosh<sup>21</sup>, Kun Qin<sup>6,7</sup>, Susan Rotzinger<sup>20,22</sup>, Matthew D. Sacchet<sup>23</sup>, Jonathan Savitz<sup>24</sup>, Haochang Shou<sup>3,25</sup>, Ashish Singh<sup>3</sup>, Aleks Stolicyn<sup>21</sup>, Irina Strigo<sup>26</sup>, Stephen C. Strother<sup>5,27</sup>, Duygu Tosun<sup>26</sup>, Teresa A. Victor<sup>24</sup>, Dongtao Wei<sup>28</sup>, Toby Wise<sup>29</sup>, Rachel D. Woodham<sup>1</sup>, Roland Zahn<sup>2</sup>, Ian M. Anderson<sup>30</sup>, J. F. William Deakin<sup>30</sup>, Boadie W. Dunlop<sup>31</sup>, Rebecca Elliott<sup>30</sup>, Qiyong Gong<sup>6,7</sup>, Ian H. Gotlib<sup>32</sup>, Catherine J. Harmer<sup>16</sup>, Sidney H. Kennedy<sup>20,22,33</sup>, Gitte M. Knudsen<sup>12,13</sup>, Helen S. Mayberg<sup>8</sup>, Martin P. Paulus<sup>24</sup>, Jiang Qiu<sup>28</sup>, Madhukar H. Trivedi<sup>9</sup>, Heather C. Whalley<sup>21</sup>, Chao-Gan Yan<sup>34</sup>, Allan H. Young<sup>2,35</sup> and Christos Davatzikos<sup>3</sup>

## Abstract

**Background** Efforts to develop neuroimaging-based biomarkers in major depressive disorder (MDD), at the individual level, have been limited to date. As diagnostic criteria are currently symptom-based, MDD is conceptualized as a disorder rather than a disease with a known etiology; further, neural measures are often confounded by medication status and heterogeneous symptom states.

**Methods** We describe a consortium to quantify neuroanatomical and neurofunctional heterogeneity via the dimensions of novel multivariate coordinate system (COORDINATE-MDD). Utilizing imaging harmonization and machine learning methods in a large cohort of medication-free, deeply phenotyped MDD participants, patterns of brain alteration are defined in replicable and neurobiologically-based dimensions and offer the potential to predict treatment response at the individual level.

International datasets are being shared from multi-ethnic community populations, first episode and recurrent MDD, which are medication-free, in a current depressive episode with prospective longitudinal treatment outcomes and in remission. Neuroimaging data consist of de-identified, individual, structural MRI and resting-state functional MRI with additional positron emission tomography (PET) data at specific sites. State-of-the-art analytic methods include

\*Correspondence:

Cynthia H. Y. Fu  
c.fu@uel.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

automated image processing for extraction of anatomical and functional imaging variables, statistical harmonization of imaging variables to account for site and scanner variations, and semi-supervised machine learning methods that identify dominant patterns associated with MDD from neural structure and function in healthy participants.

**Results** We are applying an iterative process by defining the neural dimensions that characterise deeply phenotyped samples and then testing the dimensions in novel samples to assess specificity and reliability. Crucially, we aim to use machine learning methods to identify novel predictors of treatment response based on prospective longitudinal treatment outcome data, and we can externally validate the dimensions in fully independent sites.

**Conclusion** We describe the consortium, imaging protocols and analytics using preliminary results. Our findings thus far demonstrate how datasets across many sites can be harmonized and constructively pooled to enable execution of this large-scale project.

**Keywords** Classification, Biomarkers, Deep learning, Neuroimaging, Depression, Harmonization, Predictors, MRI

## Background

Depression has been recognized for millennia as a distinct illness, included in what Hippocrates termed ‘melancholia’ and posited to be caused by black bile. Current diagnostic criteria for major depressive disorder (MDD) are based solely on a set of symptoms and observable behaviors. MDD is characterised by a persistent low mood and/or an inability to experience usual feelings of enjoyment, associated with disturbances in sleep, appetite and psychomotor functioning, low energy, poor concentration, guilt or worthlessness, and, for some, suicidal ideation and behaviors [1, 2]. MDD is highly prevalent and has significant personal, familial and socioeconomic impacts [3, 4]. At the present time, MDD remains a syndrome without an identified etiology, rather than a disease with a demonstrable pathology. There are no neurobiological markers that can identify the diseases which comprise a clinical MDD diagnosis. We lack reproducible neurobiological markers to improve the etiological and prognostic homogeneity as well as to predict response to treatment.

Not only can heterogeneous combinations of symptoms fulfil diagnostic criteria, but current criteria do not fully capture the range of symptoms. For example, high levels of anxiety and comorbid anxiety disorders are present in 50–75% of MDD individuals, which are further linked with impaired treatment response and chronic longitudinal course [5, 6]. Heterogeneity among clinical profiles can lead to less predictable responses to a given treatment, an inability to predict the longitudinal course for individual patients, and the symptom variability in ‘gold standard’ scales designated to evaluate efficacy of therapeutic interventions. Further, treatment outcomes are frequently unsatisfactory. For 30–40% of MDD individuals in a current depressive episode, an adequate treatment response or remission is not achieved even after several trials of pharmacotherapy or psychotherapy over a year [7, 8].

It is unlikely that MDD is caused by a single etiological factor. MDD heritability estimates range from 28 to 44% [9] and are considerably lower than estimates for bipolar disorder and schizophrenia, which range from 60 to 90% [10, 11]. Genome-wide association and candidate gene studies have powerfully demonstrated polygenic heritability, consisting of hundreds of variants and genes with each having a small genetic contribution. However, genetic risk variants have not clinically useful at the individual level for diagnosis, and it is unclear how genetic risk ultimately translates into an acute depressive episode [12].

Neuroimaging-based biomarkers can help to identify the various disease components that comprise MDD. Genetic and environmental factors that lead to MDD are expressed as subtle and widespread alterations in brain structure and brain function. Research over two decades provides convincing evidence of morphometric and neural alterations in MDD, despite limitations in diagnosis and treatment selection. An overarching aim has been to delineate the neurobiological features that comprise MDD and to develop imaging markers in this disorder [13–15]. It is important to recognize though that to search for a neural signature that wholly replicates current diagnostic criteria would be circular, *petitio principii*, because current diagnostic criteria are based on clinical features and different mechanisms could lead to the same clinical presentation. Efforts have been further hindered by neural measures which have been confounded by a mixture of depressive states, multiple and longstanding treatments, as well as comorbid disorders.

In recent years, machine learning based methods for MRI heterogeneity analysis have been developed which detect and characterize neuroanatomical heterogeneity of disease using a data-driven approach that generates quantifiable, replicable and neurobiologically-based metrics of disease subtypes [16]. Large samples have recently been created from multi-site datasets, however the multivariate pattern analysis has been limited to either resting-state

functional MRI data or structural MRI data with limited clinical phenotyping and lack of longitudinal treatment outcome data. From resting-state functional MRI data, four functional connectivity patterns were observed in frontostriatal and limbic systems which reflected different symptom profiles [17]. However, the training and cross-validation samples were confounded by medication and form of depression, namely treatment-resistant depression, which has demonstrable effects on brain structure and function [18–20], and there was no independent testing in a novel sample. Moreover, the models were not reproduced in an independent study [21], which could reflect overfitting in the clustering algorithm as well as insufficient overlap in the sample characteristics. Resting state functional MRI reflects depressive state, and these samples had distinct depressive states, multiple comorbid disorders, different forms of depression, with neurofunctional correlates that were likely underpowered for such non-overlapping samples [17–21].

In structural MRI, recent multi-site cohorts show classification accuracies ranging from 52 to 75% [22–25]. However, all the dimensions have been binary (ie. either MDD or healthy control). The highest classification accuracy was achieved in a cohort with a MDD diagnosis based on diagnostic criteria and that was in a current depressive episode, but the sample size was limited (230 MDD, 77 controls) [22]. In the largest sample to date (2288 MDD and 3077 controls), the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium found a classification accuracy up to 62% [23]. However, the ENIGMA MDD sample consists of a wide range of clinical phenotypes, with multiple comorbid disorders, many forms of depression from first episode to treatment-resistant depression, but limited medication history, antidepressant dosage or duration, and no treatment outcome data [26]. Similarly, large samples from UK Biobank data have limited treatment history, no treatment outcome data, and are based on a probable diagnosis of lifetime MDD derived from self-report symptoms in a population-based sample [27].

Our consortium aims to identify imaging signatures of disease heterogeneity in MDD using structural and resting state functional MRI. This will generate a neuroanatomical-neurofunctional (NA-NF) dimensional coordinate system (COORDINATE-MDD) in which each dimension captures a distinct pattern of brain alterations. Our aim is to identify the multivariate dimensions that define disease-related phenotypes in MDD and the distinct dimensions that predict treatment response at the individual level. Importantly, this is an iterative process to: (i) define neural dimensions in deeply phenotyped participants who are medication-free and in a current depressive episode in order to

delineate state and trait status; and (ii) test dimensions in novel samples to assess specificity and reproducibility. With this aim, our consortium combines extensive datasets of ‘raw’ individual-level neuroimaging and deeply phenotyped clinical data, using state-of-the-art analytic methods for big data and semi-supervised clustering. The present sample consists of richly phenotyped, individual-level data from participants with first episode or recurrent MDD, that is not treatment-resistant depression, antidepressant medication-free, with prospective longitudinal treatment outcomes, and healthy controls. The current focus is on structural and resting-state functional MRI.

We have sought to focus on first episode and recurrent MDD in the present sample. Treatment-resistant depression is currently a clinical criterion that refers to a form of depression which shows significant persistent symptoms despite a series of treatments. If treatment resistance is present at the first episode, then it might be possible to identify this dimension early in the illness. Low rates of remission associated with current treatments demonstrate their limitations [28] and could also indicate a subgroup that will progress to fulfil the clinical criterion of treatment resistance. It is possible that the pathophysiology of treatment-resistant depression might be characterized early in the course of illness [14].

Sample size directly influences the capacity of machine learning methods to reliably identify imaging signatures of disease from MRI data and machine learning-based signatures will need to be replicated in independent cohorts. Through international collaborations, we are bringing together a large and integrated sample. MRI images are processed using image processing methods that leverage robust and fully automated pipelines for extracting structural and functional imaging features. In large multi-site datasets, harmonization of imaging features from each site is a critical requirement. We will apply a statistical harmonization methodology, Combat-GAM [29], developed for pooling neuroimaging data across multiple scanners and cohorts with diverse age ranges and with the presence of nonlinear age-related differences in brain images. Combat-GAM has been shown to remove unwanted sources of variability, specifically site differences, while preserving variations due to biologically-relevant covariates in the data. Harmonized data from different scanners and sites are analyzed using machine learning and deep learning methods for imaging pattern analysis. These methods integrate small yet coordinated brain effects into signatures that may yield high sensitivity and specificity in characterizing disease effects in individuals.

The dimensional approach naturally extends to other modalities, including diffusion tensor imaging,

task-based functional MRI, and to molecular brain imaging with PET. We will extend these findings to studies of at-risk individuals and to clinical trials in an iterative process [30] in which the position and trajectory of a new individual in a broader dimensional system can be interrogated along with any number of clinical phenotypes, with the potential to study such cohorts over the course of illness and with treatment.

### Methods and preliminary results

To establish a NA-NF dimensional system of brain imaging biomarkers and predictors of clinical outcome, we will apply the following steps:

- 1) Inter-site harmonization and image analysis to create a methodological platform for constructive integration of structural imaging and neurofunctional connectivity data from multiple sites.
- 2) To investigate heterogeneity in neuroanatomy and functional connectivity as a collection of NA-NF patterns or dimensions, we will use semi-supervised AI methods, instead of commonly used clustering (unsupervised) approaches, in order to delineate disease-related effects, rather than variations in brain morphology and physiology that might be caused by a number of disease-irrelevant factors. Validity and relevance of identified dimensions will be verified by cross-validation and replication in independent datasets. Critically, we will derive the presence of each NA-NF signature in each participant.
- 3) To evaluate individual NA-NF signatures for their ability to predict or to moderate treatment response, two complementary approaches will be applied. The

first approach is wholly data-driven and will assess the degree of expression of each signature to evaluate whether the position in the NA-NF dimensional system predicts treatment response. The second approach incorporates knowledge of treatment type and clinical response to enrich the NA-NF system by dimensions that are strong predictors in order to assess whether accuracy improves with the interaction of NA-NF dimension and clinical knowledge of treatment type and response.

### Participating studies and datasets

An invitation for participation was made to research groups to share ‘raw’ neuroimaging data. The present consortium provided medication-free participants with first episode or recurrent MDD, that is not treatment-resistant depression, and healthy controls.

Collaborations have been established with centers worldwide from Canada, China, European Union, United Kingdom and USA. Each center has completed a Data Sharing Agreement to provide de-identified data in accordance with institutional policies and applicable federal, state or local laws and regulations, including ethical approvals. Summary descriptions are below (Tables 1 and 2):

Canadian Biomarker Integration Network in Depression (CAN-BIND) is a national depression program with recruitment from 7 centers [31, 32]. The CAN-BIND-1 treatment protocol is an 8-week trial with a selective serotonin reuptake inhibitor (SSRI) antidepressant, escitalopram, followed by an 8-week augmentation with aripiprazole if there was poor treatment response (i.e., less

**Table 1** Demographic, clinical and neuroimaging data

---

Clinical assessments
Demographic data (age, sex, ethnicity, handedness, IQ, years of education)
Standardised diagnostic criteria (DSM or ICD)
Psychiatric history (including comorbid disorders)
Form of depression (first episode, recurrent, treatment resistant)
Treatment (current, history)
Depression rating scales (HRSD, MADRS, QIDS)
Depressive severity (mild, moderate, severe)
Treatment outcome (baseline and post-treatment rating scale scores, remission, response)
Neuroimaging
Structural MRI
Resting-state fMRI
Diffusion tensor imaging
PET
EEG
Neuropsychological assessments

---

**Table 2** Preliminary study demographics, MRI field strength and MRI scanner model

Study	Number	Mean age	Age range	MRI	MRI scanner model
CAN-BIND	309	34	18–61	3T	GE Discovery MR750
				3T	GE Signa HDx
				3T	Philips Intera
				3T	Siemens TrioTim
EMBARC	336	37	18–65	3T	GE Discovery MR750
				3T	GE Signa HDx
				3T	Philips Achieva
				3T	Philips Ingenia
KCL	40	30	18–45	3T	Siemens TrioTim
				3T	GE Discovery MR750
				3T	Philips Intera
				3T	Siemens TrioTim
LIBR	296	32	18–59	3T	GE Discovery MR750
Manchester	70	35	20–56	1.5T	Philips Intera
NeuroPharm1	207	28	18–59	3T	Siemens Prisma
Oxford	70	30	19–61	3T	Siemens TrioTim
PRedICT	344	40	18–64	3T	Siemens TrioTim
HMRRC	269	31	18–60	3T	GE Excite
				3T	Siemens TrioTim
Stanford	110	33	19–58	3T	GE Signa-Excite
STRADL	1189	60	26–84	3T	Philips Achieva-TX
				3T	Philips Prisma-FIT

than 50% improvement in depressive symptoms). MRI scans were acquired at baseline, weeks 2 and 8 in both MDD and healthy participants.

Copenhagen University (NeuroPharm1) cohort consists of data from the Center for Integrated Molecular Brain Imaging (Cimbi) and Center for Experimental Medicine Neuroparmacology prospective longitudinal treatment study with an SSRI, escitalopram. The protocol is a 12-week trial with a SSRI, escitalopram, which could be switched to a serotonin norepinephrine reuptake inhibitor, duloxetine at week 4 if there were unacceptable side effects or poor treatment response (i.e., less than 25% improvement in depressive symptoms). MRI scans were acquired at baseline in healthy participants and at baseline and week 8 in MDD [33–39].

Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC) is a multisite, randomized, placebo-controlled clinical trial

with recruitment from 4 centers [40]. Treatment protocol consists of two stages. Stage 1 is an 8-week, double-blind, randomized allocation to placebo medication or to an SSRI, sertraline, followed by Stage 2, an 8-week, double-blind, cross over treatment design. At Stage 2, participants continue treatment for 8 weeks (either placebo or sertraline) if the Clinical Global Improvement scale (CGI) rating is at least “much improved”. If the CGI rating is less than “much improved”, then treatment is switched under double-blind conditions. From the initial placebo treatment arm, treatment is switched to the SSRI, sertraline, and from the initial SSRI, sertraline, treatment arm, the treatment is switched to the non-serotonergic antidepressant, bupropion. MRI scans were acquired in medication-free MDD and healthy participants at baseline.

Huaxi MR Research Center (HMRRC) cohort consists of medication-naïve first episode MDD and matched healthy participants with single session MRI scans [41–43].

King’s College London cohort consists of 4 studies [44–50]. Treatment protocol was an 8-week trial of serotonin-norepinephrine reuptake inhibitor (SNRI), duloxetine. MRI scans were acquired at baseline, weeks 2 and 8 in both MDD and healthy participants [47].

Laureate Institute for Brain Research (LIBR) cohort consists of 2 studies with MRI data in first episode and recurrent MDD and matched healthy controls [51–53].

University of Manchester cohort consists of 3 studies [54–56]. Treatment protocol was an 8-week trial of SSRI, citalopram, with MRI scans acquired at baseline and week 8 in both MDD and healthy participants [54].

University of Oxford cohort consists of an 6-week trial of SSRI, escitalopram, with MRI scans acquired at baseline and week 6 [57, 58].

Predictors of Remission in Depression to Individual and Combined Treatments (PRedICT) study is a 12-week randomized clinical trial of treatment-naïve MDD with 3 treatment arms: an SSRI, escitalopram; an SNRI, duloxetine; or CBT, and a 12-week second phase if remission was not achieved with monotherapy, with the addition of CBT to the medication treatment arms or augmentation with escitalopram to the CBT treatment arm. MRI scans were acquired at baseline and week 12 [59, 60].

REST-meta-MDD study consists of resting state fMRI data in medication-naïve first episode and recurrent MDD from 17 sites in China [61, 62].

Southwest University (SWU) cohort consists of a community-based recruitment which includes first episode and recurrent MDD and healthy control participants [63–65].

Stanford University cohort consists of MRI data in first episode and recurrent MDD and healthy control participants [66–68].



STratifying Resilience and Depression Longitudinally (STRADL) is a community-based cohort from the Generation Scotland Scottish Family Health Study with detailed clinical, cognitive and neuroimaging assessments. Single session MRI scans were acquired [24, 69].

University of California at San Francisco (UCSF) cohort consists of first episode and recurrent MDD and healthy control participants [70, 71].

### Assessments

MDD diagnosis is based on standardised diagnostic criteria, DSM or ICD: DSM-IV (HMRRRC, Manchester, Oxford, PReDICT, SWU), DSM-IV-TR (CAN-BIND, EMBARC, KCL, LIBR, Stanford, STRADL) and DSM-5 (Neuropharm1). Structured clinical interview assessments were performed: Structured Clinical Interview for DSM (SCID) (EMBARC, HMMRC, KCL, LIBR, Manchester, Oxford, PReDICT, Stanford, SWU, STRADL) or Mini International Neuropsychiatric Interview (MINI) (CAN-BIND, KCL, Neuropharm1).

Symptom severity has been measured using standardised clinician-rated symptom scales: 17-item Hamilton Rating Scale for Depression (HRSD) [72] (EMBARC, HMRRRC, KCL, LIBR, Neuropharm1, Oxford, Stanford, SWU), 24-item HRSD (PReDICT), Montgomery-Åsberg Depression Rating Scale (MADRS) [73] (CAN-BIND, LIBR, Manchester; PReDICT), and Quick Inventory for Depressive Symptomatology (QIDS-SR) [74, 75] (EMBARC, PReDICT, STRADL).

Clinical measures include illness history (e.g., age onset, number of previous episodes) and medication records. Demographic information (e.g., years of education) and neuropsychological assessments, including memory and executive function, where available are included (CAN-BIND, EMBARC, KCL, LIBR, Manchester, Neuropharm1, PReDICT, STRADL).

Prospective longitudinal treatment studies have provided clinical outcome data (CAN-BIND, EMBARC, KCL, Manchester, Neuropharm1, Oxford, PReDICT). We have additional information in MDD participants taking antidepressant medication (CAN-BIND, KCL, Manchester, Neuropharm1, Oxford, PReDICT, Stanford, STRADL, SWU).

### Clinical data harmonization

High reliability, internal consistency and correspondence have been demonstrated for HRSD [76], between HRSD and QIDS [77, 78], HRSD and MADRS [79], and QIDS and MADRS [80] in outpatient non-psychotic MDD. Standardized conversions will be applied to generate comparable scores [77, 78, 81]. Clinical remission is defined as 17-item HRSD score of less than or equal to 7 and the equivalent in MADRS and QIDS, and treatment response is defined as having an improvement of at least 50% relative to baseline depressive severity [8].

It is possible that derived NA-NF dimensions could be transdiagnostic. We will also incorporate clinical knowledge using a Research Domain Criteria (RDoC)-informed approach to specify transdiagnostic dimensions, using individual items as well as symptom clusters scores from rating scales [82–85]. For example, within the Arousal and Regulatory System domain, we can specify fine-grained disturbances in sleep (i.e., initial, middle and late insomnia, as well as hypersomnia) and vegetative disturbances (i.e., appetite increase or decrease, weight increase or decrease), and in the Negative Valence Systems domain, a dimension based on items assessing anxiety reflecting sensitivity to potential threat.

While all patients have a diagnosis based on standardised criteria, we cannot rule out the possibility of site differences reflecting true cultural or ethnic differences in MDD phenomenology. Potential systemic differences in symptom profiles across sites due to different socio-cultural contexts will be examined. Common factor structures have been demonstrated across cultures [86, 87]. Principal component analyses will be performed to examine factor structures that characterize symptom profiles by site. We will control for any site differences in the magnitude of symptoms by deriving within-center standardized scores using the residuals. If necessary, we will apply factorization methods to the standardized scores to obtain illness dimensions, which we will validate by analyzing their associations with demographic, diagnostic, neurocognitive and other clinical data. We will empirically test the brain-to-symptom mappings by altering the balance between inter-site pooling and within-site normalization.

### MRI imaging characteristics

MRI data include structural MPRAGE or equivalent structural MRI scans and resting state functional MRI scans acquired on 1.5 Tesla (T) or 3T MRI systems. Different models of scanner have been used, including Discovery MR750 3T (GE Healthcare, Little Chalfont, Buckinghamshire, UK), Signa HDxt 3T (GE Healthcare, Little Chalfont, Buckinghamshire, UK), MAGNETOM TrioTim or MAGNETOM Prisma Fit (Siemens Healthcare, Erlangen, Germany), and Achieva 3T (Philips Healthcare, Best, Netherlands) (Table 2).

### MRI data integration and harmonization

Harmonization is a critical process as it enables the constructive pooling and integration of all datasets within the consortium. Although imaging protocols among all included studies are comparable, scanner and minor acquisition protocol variations nonetheless introduce inter-study differences in imaging characteristics which

render direct pooling of the data very difficult. In the present consortium, we have harmonized image processing pipelines, statistical harmonization of derived measures, and deep learning based harmonization of raw images [88].

#### ***Image-level extraction of structural anatomy and functional connective measures***

Imaging variables are extracted from the MRI scans to provide a multi-scale representation of structural and connectomic characteristics. Some features are defined a priori (e.g., multi atlas-based parcellations) and some are data-driven (e.g., structural covariance and functional connectivity networks).

***MUSE region of interest segmentation*** Volumetric MRI features are extracted via established and validated methods using a fully automated processing pipeline. Raw 3D T1-weighted MRIs are first quality checked for motion, image artifacts, or restricted field-of-view. Each participant's quality-controlled T1-weighted MRI scan is preprocessed with a containerized processing pipeline. Preprocessing steps consist of magnetic field intensity inhomogeneity correction [89] and multi-atlas skull-stripping for the removal of extra-cranial material [90].

The images are segmented using a state-of-the-art, multi-atlas, multi-warp label-fusion method, MUSE [91]. In this framework, multiple atlases with semi-automatically extracted ground-truth region of interest (ROI) labels are first warped individually to the target image using two different non-linear registration methods. A spatially adaptive weighted voting strategy is then applied to fuse the ensemble into a final segmentation. This procedure is used to segment each image into 145 anatomical regions of interest (ROIs) spanning the entire brain. We calculate the volumes of the 145 ROIs, as well as the volumes of 113 composite ROIs that are obtained by combining individual ROIs into larger anatomical regions following a predefined ROI hierarchy whereby brain anatomy can be quantified at multiple levels of resolutions [91].

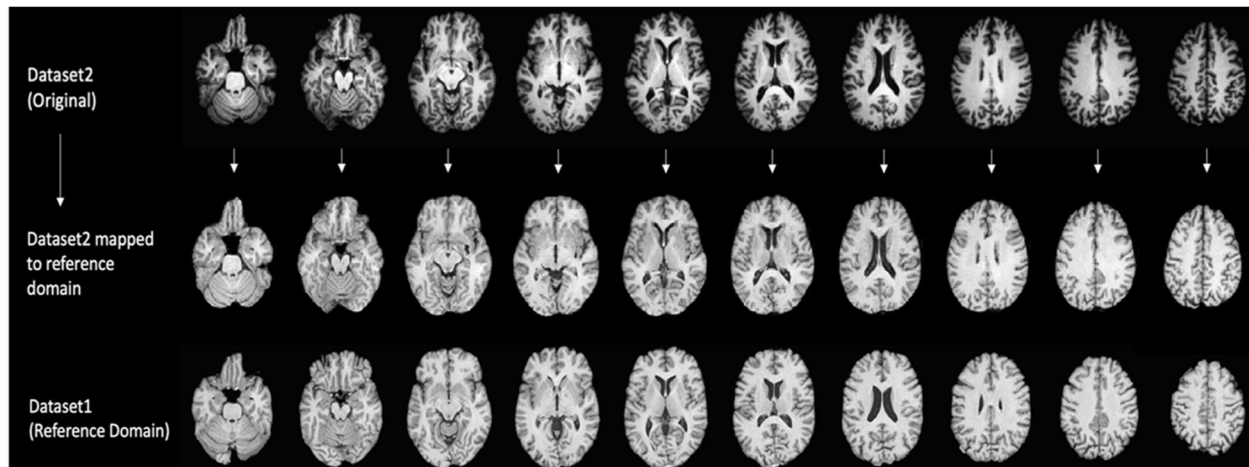
MUSE has obtained top accuracy in comparison with multiple benchmark methods in independent evaluations [92]. MUSE utilizes state-of-the-art multi-atlas, multi-warp algorithms, with a very rich set of atlases spanning several protocols from 1.5T SPGR images to 3T MPRAGE images. The ensemble of atlases used to segment each scan contains sufficient diversity which renders it robust to scanner variations [91]. Critically, we have leveraged several studies which have scanned the same individuals in multiple scanners and hence have

built procedures that render the multi-scanner atlases mutually consistent [91].

Voxel-based, tissue-specific (grey matter, white matter, cerebrospinal fluid) regional volumetric maps are obtained using MUSE segmentation and inverse atlas warping is applied for voxel-based analyses. Voxel-wise regional volumetric maps (RAVENS) are generated for each tissue [93] by spatially aligning the skull-stripped images to a template in MNI-space [94]. Quality control (QC) on derived maps and imaging variables is performed using a semi-automated approach. An automated procedure automatically ranks scans based on outlying values of quantified metrics (i.e., ROI values) and flags those that show deviation from estimated population distributions. Flagged images are examined by visual inspection using a visualization and annotation tool for evaluating for pipeline failures (e.g., poor brain extraction, tissue segmentation, and registration errors).

***Structural covariance networks*** Our prior work has highlighted the potential of structural covariance networks as a flexible and biologically interpretable way of reducing complex anatomical images down to a relatively small set of measurements [95, 96]. This approach uses orthogonally-projective non-negative matrix factorizations to parcellate the brain into regions that show consistent trends across individuals, potentially because they are influenced by common underlying neurobiological factors. For example, we found that patterns of structural covariance were highly reproducible, aligned well with functional networks, displayed differential developmental trajectories during adolescence, and correlated with maps of evolutionary cortical expansion [96]. We have recently developed a novel method that computes multi-scale non-negative matrix factorization (NMF) components [97]. This allows us to characterize heterogeneous presentation of neuropsychiatric and neurodegenerative diseases in multiple scales. These networks will be part of our imaging feature panel and might be well-suited for capturing the neurodevelopmental underpinnings of MDD.

***Deep learning generative adversarial networks (GAN)*** A second approach utilizes deep learning CycleGAN and STAR-GAN methods [88], which synthesize images drawn from certain distributions. This method has been used extensively in numerous applications seeking to transform images from one style to another. We have used data from numerous scanners to learn the mapping of brain MRI scans to a canonical

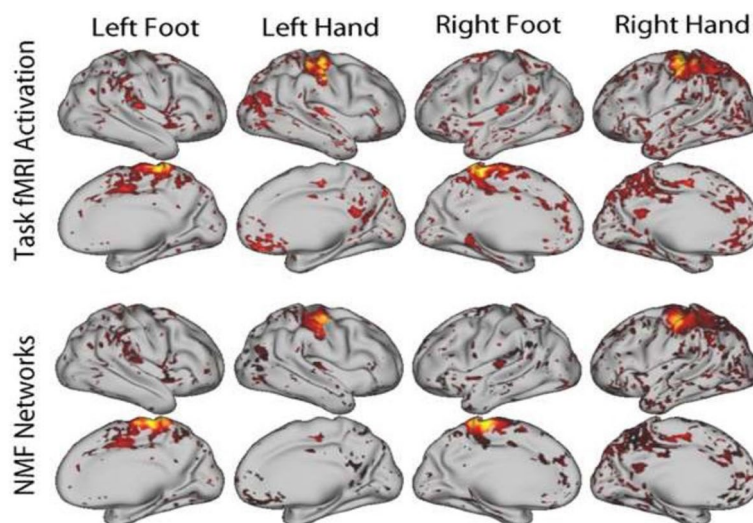


**Fig. 1** Example of image harmonization via deep learning based canonical mapping to a reference domain presented in bottom row

reference domain, in which inter-scanner variations are minimized [88] (Fig. 1). By making the characteristics of images very similar, while preserving important information, subsequent image analysis steps, including segmentation and parcellation, become significantly more consistent across studies. We will further refine and adapt this method to this project, including some of the current directions that involve nonlinear modelling co-variables in order to better restrict these mappings to differences related to MRI scanners and protocols, while preserving anatomical characteristics.

*Personalized functional networks* For resting state fMRI, personalized functional networks using a spatially-regularized NMF method align well with functional

activation patterns and have improved functional homogeneity [98] (Fig. 2). Functional homogeneity quantifies the degree to which a functional network represents one signal, rather than mixing multiple neurophysiological signals. Measures of functional connectivity, derived from brain parcellations and functional networks show promise in predicting brain maturity and distinguishing disease from healthy brain states [99–102]. Furthermore, there are significant correlations between elements of the connectivity matrices, which may lead to unreliable classification [103]. Using our NMF method, we found increased homogeneity compared to standard network definitions using a group atlas, as well as a null distribution created using a conservative spin-based spatial permutation procedure [104]. The personalized functional

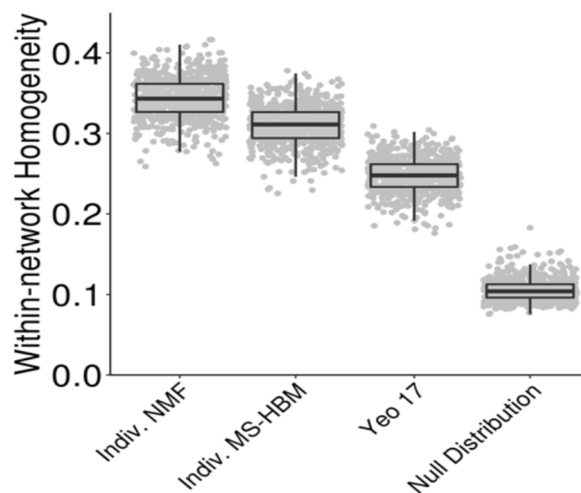


**Fig. 2** Personalized functional networks defined using NMF with resting state fMRI data aligned with activation during a motor task, adapted from Li et al. [98]



networks delineated from our NMF method showed that the networks with the greatest variability in functional topography in youth are the very higher-order association networks impacted by psychopathology, these association networks are refined with age in development, and individual differences in association network topography predicted executive function [104] (Fig. 3).

We are also employing a complementary method based on sparse connectivity pattern (SCP) learning, leveraging the effective non-linearity of sparse dictionary learning [104, 106–108] as a means for describing the functional connectivity patterns of brain networks [109]. SCPs also minimize the negative impact of correlated features on



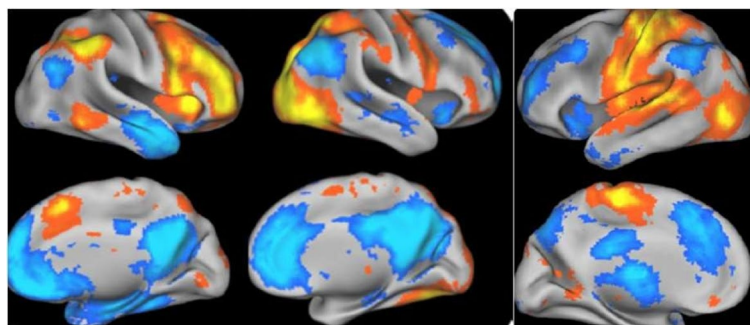
**Fig. 3** Personalized functional networks improve functional homogeneity. Functional homogeneity is higher within personalized functional networks (from either non-negative matrix factorization (NMF) or multi-session hierarchical Bayesian model (MS-HBM)) than standard group-level networks (Yeo 17-network group atlas (Yeo 17)) or a null model that preserves spatial covariance structure, adapted from Cui et al. [105]

the robustness of prediction models [103] (Fig. 4). A hierarchical extension of this approach extracts these functional networks at multiple scales [110]. Adopting these complementary techniques will provide a rich set of imaging features to characterize the functional brain connectome.

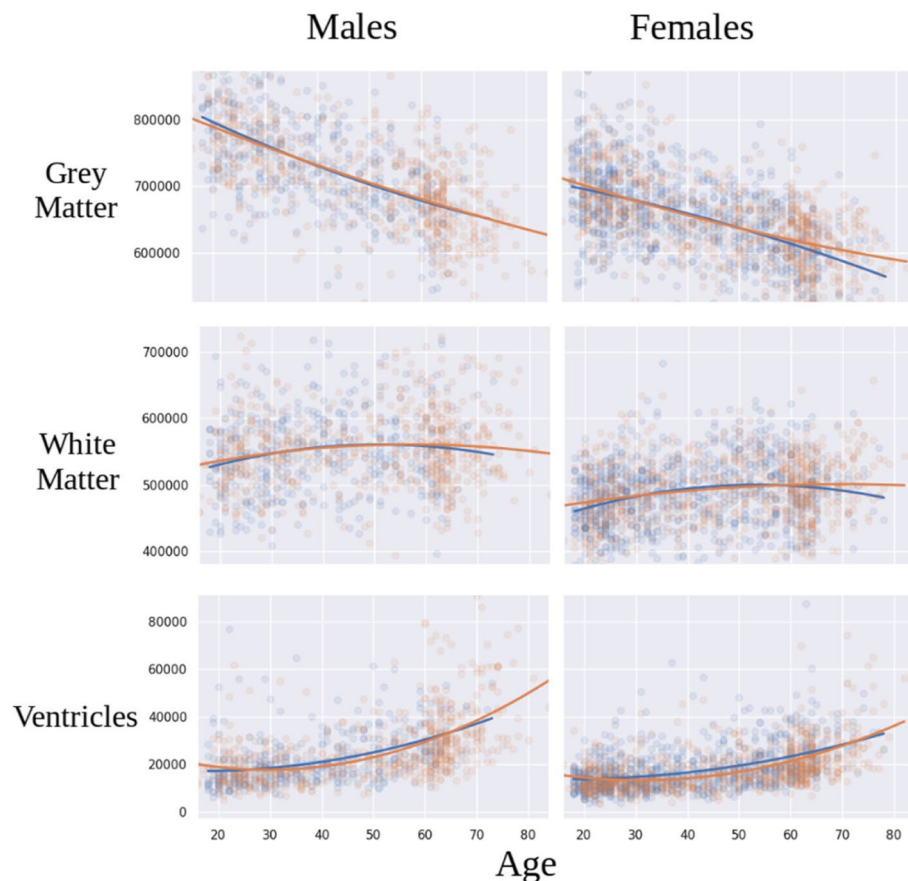
#### Harmonization of derived measures

We have developed a statistical harmonization approach [29], building upon the COMBAT method that has been successfully used for over a decade to remove batch effects in genomic studies and recently adapted in neuroimaging research [111–114]. This approach is fully multivariate, utilizing hyper-parameters to define over-arching statistical priors, and has been successfully adopted in imaging. In order to model nonlinear effects of covariates (e.g., age), we have combined COMBAT with generalized additive models (GAMs) using spline functions. The resultant COMBAT-GAM general tool for harmonization can be applied to various forms of data, including ROIs and coefficients of structural covariance and functional connectivity networks. Preliminary results of statistical harmonization from the present consortium are shown in Fig. 5.

Harmonization of resting state fMRI data is typically performed at the correlation matrix level. In particular, we have developed a functional connectivity covariance batch effect correction (FC-CovBAT) [115] that models second order moments of the upper-triangular elements of individual correlation matrices derived from the fMRI data. FC-CovBAT is an extension of COMBAT and CovBat methods [116] for structural imaging data. These methods statistically model the site/scanner differences not only in the means and variances of the multivariate correlation values, but also in the covariance structures



**Fig. 4** Three representative SCPs identified. Left column: Default mode anti-correlated with fronto-parietal network; Middle column: visual and default mode anti-correlated regions; Right column: sensorimotor regions anti-correlated with fronto-parietal network; blue and orange colours represent anti-correlated regions, adapted from Eavani et al. [109]



**Fig. 5** Preliminary results from the COORDINATE MDD datasets (EMBARC, Oxford, HMRC, Stanford, STRADL) showing age trajectories in grey matter, white matter and ventricular volumes in MDD patients (colored blue) as compared to healthy controls (colored red)

between the multivariate correlation values from the FC data (Fig. 6).

#### NA-NF dimensional neurobiological representation of MDD

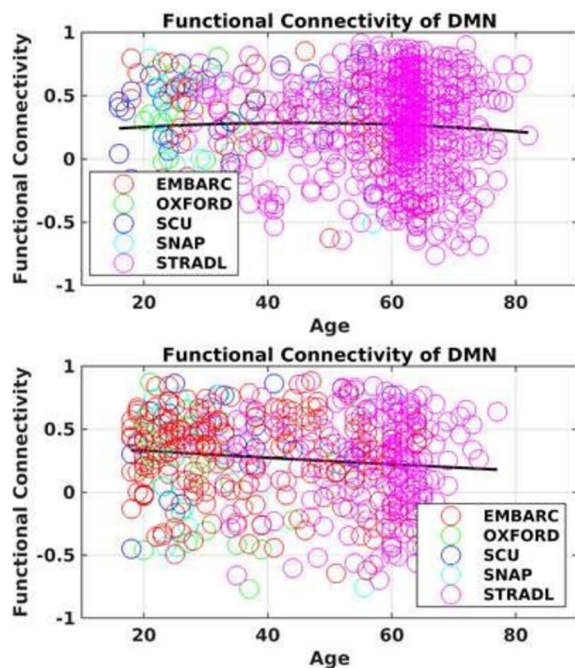
Direct clustering of patient data, especially high-dimensional images, is challenged by effects of confounding variations in the data, namely variations in healthy individuals that are not related to disease effects. We have developed semi-supervised clustering methods to account for these effects, effectively clustering differences between patients and healthy controls, rather than directly clustering patient data. The methods assume that patient data have been derived from healthy control data via one of several, to be estimated, transformations that reflect disease effects (Fig. 7). Whatever variations are present in the healthy control data and are unrelated to disease effects will follow the disease-specific transformations. Within this approach, we acknowledge that some control samples might reflect a “super-healthy” control group without any chronic diseases that may represent

resilience [117], though healthy control participants have also been recruited from the general community in the present consortium.

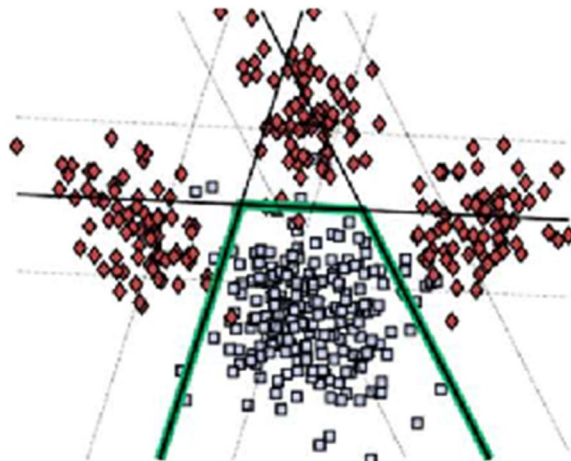
Two semi-supervised methods will be used: 1) HYDRA, a conventional machine learning approach that is largely a discriminative method [16], with its multi-scale extension [118]; and 2) Generative Adversarial Networks (GANs), a complementary method which leverages the power of a class of generative deep learning methods [119]. These methods have complementary strengths, which will be combined in seeking robust and reproducible MDD dimensions.

#### HYDRA

HYDRA uses the concept of a convex polytope, in conjunction with support vector machine principles (maximum margin cost function), to simultaneously classify patients from healthy controls and to determine a number of hyperplanes that represent disease dimensions. An iterative process determines the disease dimensions, as well as memberships in them. HYDRA has recently



**Fig. 6** Preliminary results showing functional connectivity of DMN (between anterior cingulate cortex and posterior cingulate cortex) in healthy controls (top panel) and MDD participants (bottom panel) from the COORDINATE MDD datasets (EMBARC, Oxford, SCU (HMRR), SNAP (Stanford), STRADL)



**Fig. 7** Healthy controls (green) are separated from patients via a number of maximum-margin hyperplanes that define disease subtypes. Iterative classification and clustering determine subtype membership and classification hyperplanes

helped discover two distinct neuroanatomical dimensions in schizophrenia with evidence of differences in cognitive and clinical profiles between dimensions, in

polygenic risk scores for schizophrenia and autism, as well as in brain development and aging studies [120]. We will apply this method to the imaging feature panels (ROI volumes, structural covariance network coefficients derived from voxel-based maps, and functional connectivity network coefficients), in order to identify MDD dimensions and to obtain positions of each patient within each dimension. It is likely that some patients will express multiple NA-NF signatures concurrently.

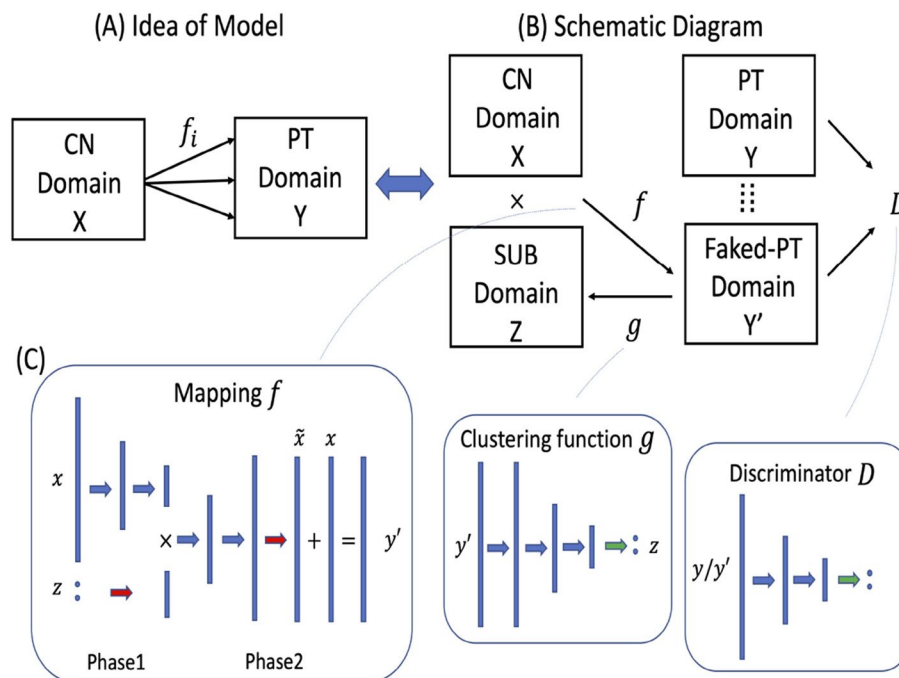
#### *Deep learning and generative adversarial networks (GANs)*

The second approach relies on a generative methodology for semi-supervised clustering, termed Smile-GAN, which utilizes state-of-the-art deep learning CycleGAN architectures, with latent variables representing neurobiological dimensions [119]. The model simultaneously learns the mapping and clustering functions using the discriminator function and the data from the patient domain. Representations of the patient clustering structure and the depression-related NA-NF dimensions are stored respectively in the network weights of functions after learning, which are the main outcomes of the model training (Fig. 8).

While these approaches can be very powerful, they require large training datasets. The availability of a sufficiently large and well-harmonized dataset in the present consortium will allow us to harness the power of these deep learning methods. We will seek experimental confirmation via split sample and sensitivity analyses.

We use cross-validation, split-sample, and random permutation experiments to determine the optimal number of reproducible and statistically significant clusters corresponding to NA-NF dimensions of MDD, by measuring the adjusted rand index (ARI) [120]. The most reproducible clusters, according to two main criteria, both high and statistically significant ARI, will be used to derive the NA-NF dimensions. As part of the method, cluster memberships of each individual will be readily calculated [16, 119]. These memberships will effectively be the coordinates of an individual in the NA-NF coordinate system, reflecting the dominant NA-NF MDD patterns.

*Incorporating clinical knowledge* We will augment this approach with the use of clinical measures, namely symptom profile as well as depressive state and medication status. The main idea is that NA-NF dimensions might not necessarily be easily interpretable clinically. By incorporating clinical phenotypic measures in parts of the clustering procedure, we will favor imaging patterns that simultaneously have relatively distinct clinical phenotypes. Methodological extensions will include, for example, introducing a discriminant term, which encodes the dimensions. This term will encourage NA-NF dimensions



**Fig. 8** Our deep learning-based approach to subtyping. **A** General idea behind Smile-GANs **(B)** Schematic diagram of Smile-GANs. CN: healthy control, PT: patient, SUB: subtype **(C)** Network architectures of the three main functions ( $f$ ,  $g$ ,  $D$ ): blue arrow represents one linear transformation followed by one leaky relu function, green arrow represents one linear transformation followed by one softmax function, red arrow represents only one linear transformation. The GAN network learns to synthesize MDD patient scans from scans of normal controls, contingent upon subtype, SUB, which is learned in the training process. The discriminator,  $D$ , of this GAN ensures that these synthesized scans are indistinguishable from real MDD patients scans. The mappings,  $f$ , estimated in this process capture features of the neuropathologic processes that transform imaging data of controls to those of MDD patients, adapted from Yang et al. [119]

that have relatively distinct clinical phenotypes. We will also modify the latent space formulation to allow for continuous latent variables, in order to better model anatomy and function in individuals who might present mixed patterns.

#### Multidimensional predictors of treatment responses

A main goal is to investigate whether position in the NA-NF dimensional system captures disease effect in predicting treatment response at the individual level based on semi-supervised approaches. Treatment response will be defined as the harmonized change in clinician-rated depressive severity scales. Changes of scores after the treatment period will be modelled both as continuous values and binary variables. We recognise that these measures are often insufficient for capturing treatment efficacy and might not reflect patient experience. We will further investigate whether NA-NF dimensions might map onto subsets of the scales with greater clinical properties [121].

We will further enrich NA-NF dimensions by directly and specifically looking for patterns that predict

treatment response via supervised classification. Such imaging signatures might or might not capture broad and treatment-oblivious NA-NF variations, however they capture NA-NF measures that are tailored to prediction of treatment response in which baseline clinical variables will be part of these treatment-specific predictive models.

In particular, specificity to particular categories of treatment will be investigated. Prediction of differential response between placebo and antidepressant medication is highly important, yet difficult to assess. Prediction of placebo response is uniquely accessible in the present sample, in which treatment outcome data are available for both active antidepressant and placebo arms of randomized controlled trials.

To identify pre-treatment moderators of differential treatment response, the specificity of NA-NF predictors to different categories of antidepressant medication will be investigated. Treatment outcome data are available for SSRI and SNRI medication classes. Specificity of NA-NF predictors of response to psychotherapy or to antidepressant medication will also be investigated.

A key weakness of multivariate pattern analysis studies has been the frequent lack of generalizability and



validation on new datasets. Given the size of the consortium, we will test the NA-NF dimensional system in wholly independent prospective longitudinal treatment cohorts, namely Copenhagen University. We emphasize that cross-validation methods are engrained in all our analyses to reduce overfitting and improve generalizability.

## Summary

COORDINATE-MDD brings together deeply phenotyped clinical data, multi-site 'raw' individual-level structural and functional neuroimaging data, and state-of-the-art AI-based methods in order to identify the multimodal dimensions that comprise MDD and which inform treatment response. Our consortium data are derived from adults with first episode or recurrent MDD, medication-free, in a current major depressive episode, that is not treatment-resistant depression, and healthy controls. We leverage statistical harmonization, machine learning and deep learning methods and large integrated and harmonized sample of highly-phenotyped, individual-level patient data with prospective longitudinal treatment outcomes. Our objectives are to delineate the robust and reproducible neurobiological NA-NF signatures which comprise MDD and which predict or moderate treatment response. We will follow an iterative process by first defining the multivariate neural dimensions that characterise deeply phenotyped samples and then testing the utility of these dimensions in novel samples.

## Abbreviations

MDD	Major depressive disorder
NA-NF	Neuroanatomical-neurofunctional
SSRI	Selective serotonin reuptake inhibitor
CAN-BIND	Canadian Biomarker Integration Network in Depression
Cimbi	Center for Integrated Molecular Brain Imaging
EMBARC	Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care
CGI	Clinical Global Improvement scale
HMRRRC	Huaxi MR Research Center
SNRI	Serotonin-norepinephrine reuptake inhibitor
LIBR	Laureate Institute for Brain Research
PReDICT	Predictors of Remission in Depression to Individual and Combined Treatments
SWU	Southwest University
STRADL	STratifying Resilience and Depression Longitudinally
UCSF	University of California at San Francisco
MINI	Mini International Neuropsychiatric Interview
HRSD	Hamilton Rating Scale for Depression
MADRS	Montgomery-Åsberg Depression Rating Scale
QIDS-SR	Quick Inventory for Depressive Symptomatology
RDoC	Research Domain Criteria
T	Tesla
ROIs	Regions of interest
GAN	Generative adversarial networks
NMF	Non-negative matrix factorization
SCP	Sparse connectivity pattern

GAMS	Generalized additive models
FC-CovBAT	Functional connectivity covariance batch effect correction
GANs	Generative Adversarial Networks
ARI	Adjusted rand index

## Acknowledgements

MT, CCF, CHYF and all authors wish to thank the participants, families, staff, and colleagues who made this project possible. BWD acknowledges W. Edward Craighead, PhD. The views expressed are those of the author(s) and not necessarily those of the funding agencies.

## Authors' contributions

Authors made substantial contributions to the conception of the work (CHYF, CD), design (all authors), data acquisition (all authors), analysis (CD, CHYF, GE, YF, MA, JG, HS). All authors have contributed to drafting and revising the work. All authors have approved the submitted version. All authors agree both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

## Funding

IMA received funding from the Medical Research Council (MRC) (G0601526). DA received funding from the Academy of Medical Science (AMS-SGCL8) and National Institute for Health and Care Research (NIHR) PhD studentship. BNF, SHK, SH, SCS, SRA, and SR received funding from Ontario Brain Institute and Canadian Institutes of Health Research (CIHR) (for CANBIND01 data). CD received funding from the National Institute of Health (R01 MH112070), HC (RF1-AG054409, R01-MH123550, U01-AG068057), IHG (R37MH101495) and YF (R01 AG066650 and R01EB022573). EMBARC study (NCT01407094) was supported by the National Institute of Mental Health of the National Institutes of Health under award numbers U01MH092221 (Trivedi, M.H.) and U01MH092250 (McGrath, P.J., Parsey, R.V., Weissman, M.M.), and in part by the Hersh Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Valeant Pharmaceuticals donated the Wellbutrin XL used in the study. This work was supported by the EMBARC National Coordinating Center at UT Southwestern Medical Center, Madhukar H. Trivedi, M.D., Coordinating PI, and the Data Center at Columbia and Stony Brook Universities. In addition, this work was funded in part by the Center for Depression Research and Clinical Care (PI: Madhukar Trivedi). RE received funding from the National Institute for Health and Care Research (NIHR), and the Medical Research Council (MRC). CHYF received funding from the Medical Research Council (MRC) (G0802594) and Brain and Behaviour NARSAD Young Investigators Award. QG received funding from the National Natural Science Foundation of China (81820108018; 81621003). CJH received support from the Medical Research Council (G0701421) and is supported by the Oxford Health NIHR Biomedical Research Centre. GMK, MG & VGF received funding from the Lundbeck Foundation (R279–2018-1145 (BrainDrugs)). HSM & BWD received funding from the National Institute of Mental Health (P50MH077083, 1R01MH080880). MPP received funding from The William K. Warren Foundation, National Institute on Drug Abuse (U01 DA041089), and National Institute of General Medical Sciences Center Grant Award Number (1P20GM121312). MDS received funding from the National Institute of Mental Health (R01MH125850). JS received funding from the National Institute of Mental Health (K01MH096077; R01MH098099). STRADL study was supported and funded by the Wellcome Trust Strategic Award, Stratifying Resilience and Depression Longitudinally (ref. 104036/Z/14/Z), and the Medical Research Council (MRC-MC/PC/17209). AMM received funding from the Wellcome Trust (220857/Z/20/Z, 216767/Z/19/Z). AS was funded as part of the STRADL study and indirectly through the Lister Institute of Preventive Medicine award ref. 173096. Data processing used the resources provided by the Edinburgh Compute and Data Facility (ECDF) (<http://www.ecdf.ed.ac.uk/>). SWU dataset was supported by the following: National Natural Science Foundation of China (31571137; 31500885), the National Outstanding

young people plan, the Program for the Top Young Talents by Chongqing, the Fundamental Research Funds for the Central Universities (SWU1509383; SWU1509451; SWU1609177), Natural Science Foundation of Chongqing (cstc2015jcyjA10106), the Fok Ying Tung Education Foundation (151023) to JQ and DTW.

DT received funding from the National Institute of Mental Health (5R01MH101472).

TW received funding from the Anthony and Elizabeth Mellows Charitable Foundation.

CGY was supported by the National Natural Science Foundation of China (82122035) and Beijing Nova Program of Science and Technology (Z191100001119104).

AHY independent research is funded by the National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London.

RZ received funding from the Medical Research Council (MRC) (MR/T017538/1).

### Availability of data and materials

Original data can be requested from each consortium member. Derived data in the present consortium can be made available from the corresponding author on reasonable request and with permission of each consortium member.

### Declarations

#### Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations. All participating sites had obtained permission from their local Institutional Review Boards or Ethics Committees, and all participants had provided written informed consent. Research analyses specific to this study only used de-identified data. Data Sharing Agreements were reviewed and approved by the Grants Department and Contracts Offices for each of the following research centers: Ontario Brain Institute (OBI) on behalf of the Canadian Biomarker Integration Network in Depression (CAN-BIND); University of East London (UEL); University of Texas Southwestern Medical Center on behalf of the Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care study (EMBARC); Emory University on behalf of the Predictors of Remission in Depression to Individual and Combined Treatments study (PReDICT); Huaxi MR Research Center (HMRRC), Sichuan University; King's College London (KCL); Laureate Institute for Brain Research (LIBR); University of Manchester; University of Oxford; University of Pennsylvania; REST-meta-MDD consortium and the International Big-Data Center for Depression Research, Institute of Psychology, Chinese Academy of Sciences; Southwest University (SWU); Stanford University; and University of Edinburgh on behalf of STrategy Resilience and Depression Longitudinally study (STRADL).

#### Consent for publication

Not applicable.

#### Competing interests

SRA has consulted for Indoc Research Canada.

AMM has received research support from Eli Lilly, Janssen, and The Sackler Trust. AMM has also received speaker fees from Illumina and Janssen. CHYF has received grant funding from MRC, Wellcome Trust, Brain and Behaviour NARSAD, Eli Lilly, Flow Neuroscience, Baszucki Brain Research Fund Milken Institute, Rosetrees Trust.

HS has received grant funding from NIH.

SCS received research support from Brain Canada, CIHR (Canadian Institutes of Health Research), Ontario Brain Institute and CFI (Canadian Foundation for Innovation). SCS is a founder and share holder of ADMdx, Inc.

DT has received grant funding from NIH.

RZ is a private psychiatrist service provider at The London Depression Institute and co-investigator on a Livanova-funded observational study of Vagus Nerve Stimulation for Depression. RZ has received honoraria for talks at medical symposia sponsored by Lundbeck as well as Janssen. He has collaborated with EMIS PLC and advises Depsee Ltd. He is affiliated with the D'Or Institute of Research and Education, Rio de Janeiro and advises the Sciens Institute, USA. BWD has received research support from Boehringer-Ingelheim, Compass, Pathways, NIMH, Otsuka, Usona. Honoraria for Consulting: Aya Biosciences,

Myriad Neuroscience, Otsuka, Sophren Therapeutics, Cerebral Therapeutics, Sage.

CJH serves as a consultant for P1vital, Lundbeck, Servier and Compass Pathways. She holds grant income from Zogenix and J&J.

SHK has received funding for Consulting or Speaking engagements from Abbvie, Boehringer-Ingelheim, Janssen, Lundbeck, Lundbeck Institute, Merck, Otsuka Pfizer, Sunovion and Servier. He has received Research Support from Abbott, Brain Canada, CIHR (Canadian Institutes of Health Research), Janssen, Lundbeck, Ontario Brain Institute, Otsuka, Pfizer, SPOR (Canada's Strategy for Patient-Oriented Research). He has stock/stock options in Field Trip Health. GMK has received honoraria as a speaker for Sage Biogen and as a consultant for Sanos.

HSM has received grant funding from NIH; Consulting and IP Licenses Fees from Abbott Labs.

MHT received Research Support from: NIH, PCORI, AFSP; Consultant: Alkermes Inc., Alto Neuroscience Inc., Axsome Therapeutics, Boehringer Ingelheim, GH Research, GreenLight VitalSign6 Inc., Heading Health, Inc., Janssen Pharmaceutical, Legion Health, Merck Sharp & Dohme Corp., Mind Medicine Inc., Navitor, Neurocrine Biosciences Inc., Noema Pharma AG, Orexo US Inc., Otsuka Canada Pharmaceutical Inc., Otsuka Pharmaceutical Development & Commercialization, Inc. (MDD Section Advisor), SAGE Therapeutics, Signant Health, Takeda Pharmaceuticals Inc.; Editorial compensation from Oxford University Press. AHY reports declaration of interests: Paid lectures and advisory boards for the following companies: AstraZeneca, Eli Lilly, Lundbeck, Sunovion, Servier, Livanova, Janssen, Allegan, Bionomics, Sumitomo Dainippon Pharma, COMPASS, Sage, Novartis. Consultant to Johnson & Johnson and to Livanova. Received honoraria for attending advisory boards and presenting talks at meetings organised by LivaNova. Principal Investigator in the Restore-Life VNS registry study funded by LivaNova. UK Chief Investigator for Novartis MDD study MIJ821A12201. Principal Investigator on ESKETINTRD3004: "An Open-label, Long-term, Safety and Efficacy Study of Intranasal Esketamine in Treatment-resistant Depression". Principal Investigator on "The Effects of Psilocybin on Cognitive Function in Healthy Participants". Principal Investigator on "The Safety and Efficacy of Psilocybin in Participants with Treatment-Resistant Depression (P-TRD)". No shareholdings in pharmaceutical companies. Deputy Editor of BJPsych Open. Grant funding (past and present): NIMH (USA); CIHR (Canada); NARSAD (USA); Stanley Medical Research Institute (USA); MRC (UK); Wellcome Trust (UK); Royal College of Physicians (Edin); BMA (UK); UBC-VGH Foundation (Canada); WEDC (Canada); CCS Depression Research Fund (Canada); MSFHR (Canada); NIHR (UK). Janssen (UK).

Authors AS, AS, BRG, BNF, CCF, C-GY, CD, DA, DTW, GE, HCW, IHG, IMA, IS, JFWD, JQ, JS, JG, KH, KSC, KQ, MPP, MA, MDS, MG, QG, RDW, RE, SH, SR, TAV, TC, TW, VGF and YF declare that they have no competing interests to declare.

#### Author details

<sup>1</sup> Department of Psychological Sciences, University of East London, London, UK. <sup>2</sup> Centre for Affective Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK. <sup>3</sup> Center for Biomedical Image Computing and Analytics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA. <sup>4</sup> Department of Psychiatry and Behavioral Science, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates. <sup>5</sup> Rotman Research Institute, Baycrest Centre, Toronto, Canada. <sup>6</sup> Huaxi MR Research Center, Department of Radiology, West China Hospital, Sichuan University, Chengdu, China. <sup>7</sup> Research Unit of Psychoradiology, Chinese Academy of Medical Sciences, Chengdu, China. <sup>8</sup> Nash Family Center for Advanced Circuit Therapeutics, Icahn School of Medicine at Mount Sinai, New York, USA. <sup>9</sup> Department of Psychiatry, Center for Depression Research and Clinical Care, University of Texas Southwestern Medical Center, Dallas, USA. <sup>10</sup> Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Canada. <sup>11</sup> Mood Disorders Treatment and Research Centre and Women's Health Concerns Clinic, St Joseph's Healthcare Hamilton, Hamilton, Canada. <sup>12</sup> Neurobiology Research Unit, University Hospital Rigshospitalet, Copenhagen, Denmark. <sup>13</sup> Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. <sup>14</sup> Department of Psychiatry, Psychiatric Centre Copenhagen, Copenhagen, Denmark. <sup>15</sup> Department of Computer Science, University of Copenhagen, Copenhagen, Denmark. <sup>16</sup> Department of Psychiatry, University of Oxford, Oxford, UK. <sup>17</sup> Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK. <sup>18</sup> Mathison Centre for Mental Health Research and Education, University

of Calgary, Calgary, Canada. <sup>19</sup>Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, Canada. <sup>20</sup>Department of Psychiatry, University Health Network, Toronto, Canada. <sup>21</sup>Division of Psychiatry, Royal Edinburgh Hospital, University of Edinburgh, Edinburgh, UK. <sup>22</sup>Centre for Depression and Suicide Studies, Unity Health Toronto, Toronto, Canada. <sup>23</sup>Meditation Research Program, Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, USA. <sup>24</sup>Laureate Institute for Brain Research, Tulsa, USA. <sup>25</sup>Penn Statistics in Imaging and Visualization Endeavor (PennSIVE) Center, Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, USA. <sup>26</sup>Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, USA. <sup>27</sup>Department of Medical Biophysics, University of Toronto, Toronto, Canada. <sup>28</sup>School of Psychology, Southwest University, Chongqing, China. <sup>29</sup>Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. <sup>30</sup>Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK. <sup>31</sup>Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, USA. <sup>32</sup>Department of Psychology, Stanford University, Stanford, USA. <sup>33</sup>Unity Health Toronto, Toronto, Canada. <sup>34</sup>CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing, China. <sup>35</sup>South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, London, UK.

Received: 29 April 2022 Accepted: 29 December 2022

Published online: 23 January 2023

## References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders, (DSM-5-TR<sup>®</sup>). Fifth Edition, Text Revision. Arlington: American Psychiatric Pub; 2022.
- World Health Organization. International Statistical Classification of Diseases and Related Health Problems. 11th ed; 2021. Cited 2022. Available from: <https://icd.who.int/en/>.
- Kessler RC, Bromet EJ. The epidemiology of depression across cultures. *Annu Rev Public Health*. 2013;34(1):19–38.
- Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet*. 2013;382(9904):1575–86.
- Fava M, Alpert JE, Carmin CN, Wisniewski SR, Trivedi MH, Biggs MM, et al. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR\*D. *Psychol Med*. 2004;34(7):1299–308.
- Hofmeijer-Sevink MK, Batelaan NM, van Megen HJGM, Penninx BW, Cath DC, van den Hout MA, et al. Clinical relevance of comorbidity in anxiety disorders: a report from the Netherlands study of depression and anxiety (NESDA). *J Affect Disord*. 2012;137(1–3):106–12.
- Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR\*D report. *AJP*. 2006;163(11):1905–17.
- Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR\*D: implications for clinical practice. *AJP*. 2006;163(1):28–40.
- Fernandez-Pujals AM, Adams MJ, Thomson P, McKechnie AG, Blackwood DHR, Smith BH, et al. Epidemiology and heritability of major depressive disorder, stratified by age of onset, sex, and illness course in Generation Scotland: Scottish Family Health Study (GS:SFHS). Ebmeier K, editor. *PLoS One*. 2015;10(11):e0142197.
- Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. *Science*. 2015;349(6255):1489–94.
- Johansson V, Kuja-Halkola R, Cannon TD, Hultman CM, Hedman AM. A population-based heritability estimate of bipolar disorder – in a Swedish twin sample. *Psychiatry Res*. 2019;278:180–7.
- McIntosh AM, Sullivan PF, Lewis CM. Uncovering the genetic architecture of major depression. *Neuron*. 2019;102(1):91–103.
- Arnone D, Job D, Selvaraj S, Abe O, Amico F, Cheng Y, et al. Computational meta-analysis of statistical parametric maps in major depression. *Hum Brain Mapp*. 2016;37(4):1393–404.
- Fu CHY, Fan Y, Davatzikos C. Widespread morphometric abnormalities in major depression: neuroplasticity and potential for biomarker development. *Neuroimaging Clin N Am*. 2020;30(1):85–95.
- Fu CHY, Costafreda SG. Neuroimaging-based biomarkers in psychiatry: clinical opportunities of a paradigm shift. *Can J Psychiatr*. 2013;58(9):499–508.
- Varol E, Sotiras A, Davatzikos C. HYDRA: revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *NeuroImage*. 2017;145:346–64.
- Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. 2017;23(1):28–38.
- Shah PJ, Glabus MF, Goodwin GM, Ebmeier KP. Chronic, treatment-resistant depression and right fronto-striatal atrophy. *Br J Psychiatry*. 2002;180(5):434–40.
- Guo W, Liu F, Xue Z, Gao K, Liu Z, Xiao C, et al. Abnormal resting-state cerebellar–cerebral functional connectivity in treatment-resistant depression and treatment sensitive depression. *Prog Neuro-Psychopharmacol Biol Psychiatry*. 2013;44:51–7.
- de Kwaasteniet BP, Rive MM, Ruhé HG, Schene AH, Veltman DJ, Fellinger L, et al. Decreased resting-state connectivity between neurocognitive networks in treatment resistant depression. *Front Psychiatry*. 2015;6:28.
- Dinga R, Schmaal L, Penninx BWJH, van Tol MJ, Veltman DJ, van Velzen L, et al. Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). *NeuroImage*. 2019;22:101796.
- Yang J, Zhang M, Ahn H, Zhang Q, Jin TB, Li I, et al. Development and evaluation of a multimodal marker of major depressive disorder. *Hum Brain Mapp*. 2018;39(11):4420–39.
- Belov V, Erwin-Grabner T, Gonul AS, Amod AR, Ojha A, Aleman A, et al. Multi-site benchmark classification of major depressive disorder using machine learning on cortical and subcortical measures. *arXiv*. 2022; Available from: <http://arxiv.org/abs/2206.08122>. Cited 2022 Sep 28.
- Stolicyn A, Harris MA, Shen X, Barbu MC, Adams MJ, Hawkins EL, et al. Automated classification of depression from structural brain measures across two independent community-based cohorts. *Hum Brain Mapp*. 2020;41(14):3922–37.
- Winter NR, Leenings R, Ernsting J, Sarink K, Fisch L, Emden D, et al. Quantifying deviations of brain structure and function in major depressive disorder across neuroimaging modalities. *JAMA Psychiatry*. 2022;79(9):879–88.
- Schmaal L, Pozzi EC, C Ho T, van Velzen LS, Veer IM, Opel N, et al. ENIGMA MDD: seven years of global neuroimaging studies of major depression through worldwide data sharing. *Transl Psychiatry*. 2020;10(1):172.
- Shen X, Reus LM, Cox SR, Adams MJ, Liewald DC, Bastin ME, et al. Subcortical volume and white matter integrity abnormalities in major depressive disorder: findings from UK biobank imaging data. *Sci Rep*. 2017;7(1):5547.
- Gartlehner G, Wagner G, Matyas N, Titscher V, Greimel J, Lux L, et al. Pharmacological and non-pharmacological treatments for major depressive disorder: review of systematic reviews. *BMJ Open*. 2017;7(6):e014912.
- Pomponio R, Erus G, Habes M, Doshi J, Srinivasan D, Mamourian E, et al. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage*. 2020;208:116450.
- Wen J, Fu CHY, Tosun D, Veturi Y, Yang Z, Abdulkadir A, et al. Characterizing heterogeneity in neuroimaging, cognition, clinical symptoms, and genetics among patients with late-life depression. *JAMA Psychiatry*. 2022;79(5):464–74.
- Kennedy SH, Downar J, Evans KR, Feilotter H, W. Lam RW, GM MQ, et al. The Canadian biomarker integration network in depression (CAN-BIND): advances in response prediction. *CPD*. 2012;18(36):5976–89.
- Lam RW, Milev R, Rotzinger S, Andreazza AC, Blier P, Brenner C, et al. Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *BMC Psychiatry*. 2016;16(1):105.
- Brandt IM, Köhler-Forsberg K, Ganz M, Ozenne B, Jorgensen MB, Poulsen A, et al. Reward processing in major depressive disorder and

- prediction of treatment response – NeuroPharm study. *Eur Neuropsychopharmacol.* 2021;44:23–33.
34. Dam VH, Stenbæk DS, Köhler-Forsberg K, Ip C, Ozenne B, Sahakian BJ, et al. Hot and cold cognitive disturbances in antidepressant-free patients with major depressive disorder: a NeuroPharm study. *Psychol Med.* 2021;51(14):2347–56.
  35. Fisher P, Ozenne B, Ganz M, Frokajer V, Dam V, Penninx B, et al. Emotional faces processing in major depressive disorder and prediction of antidepressant treatment response – a NeuroPharm study. *J Psychopharmacol.* In Press.
  36. Ip CT, Ganz M, Dam VH, Ozenne B, Rüesch A, Köhler-Forsberg K, et al. NeuroPharm study: EEG wakefulness regulation as a biomarker in MDD. *J Psychiatr Res.* 2021;141:57–65.
  37. Ip CT, Olbrich S, Ganz M, Ozenne B, Köhler-Forsberg K, Dam VH, et al. Pretreatment qEEG biomarkers for predicting pharmacological treatment outcome in major depressive disorder: independent validation from the NeuroPharm study. *Eur Neuropsychopharmacol.* 2021;49:101–12.
  38. Köhler-Forsberg K, Jorgensen A, Dam VH, Stenbæk DS, Fisher PM, Ip CT, et al. Predicting treatment outcome in major depressive disorder using serotonin 4 receptor PET brain imaging, functional MRI, cognitive-, EEG-based, and peripheral biomarkers: a NeuroPharm open label clinical trial protocol. *Front Psychiatry.* 2020;11:641.
  39. Larsen SV, Ozenne B, Köhler-Forsberg K, Poulsen AS, Dam VH, Svarer C, et al. The impact of hormonal contraceptive use on serotonergic neurotransmission and antidepressant treatment response: results from the NeuroPharm 1 study. *Front Endocrinol (Lausanne).* 2022;13:799675.
  40. Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): rationale and design. *J Psychiatr Res.* 2016 Jul;78:11–23.
  41. Qiu L, Xia M, Cheng B, Yuan L, Kuang W, Bi F, et al. Abnormal dynamic functional connectivity of amygdalar subregions in untreated patients with first-episode major depressive disorder. *JPN.* 2018;43(4):262–72.
  42. Zhao Y, Niu R, Lei D, Shah C, Xiao Y, Zhang W, et al. Aberrant gray matter networks in non-comorbid medication-naïve patients with major depressive disorder and those with social anxiety disorder. *Front Hum Neurosci.* 2020;14:172.
  43. Zhao Y, Zhang F, Zhang W, Chen L, Chen Z, Lui S, et al. Decoupling of gray and white matter functional networks in medication-naïve patients with major depressive disorder. *J Magn Reson Imaging.* 2021;53(3):742–52.
  44. Green S, Lambon Ralph MA, Moll J, Deakin JFW, Zahn R. Guilt-selective functional disconnection of anterior temporal and subgenual cortices in major depressive disorder. *Arch Gen Psychiatry.* 2012;69(10):1014–21.
  45. Lythe KE, Gethin JA, Workman CI, Lambon Ralph MA, Deakin JFW, Moll J, et al. Subgenual activation and the finger of blame: individual differences and depression vulnerability. *Psychol Med.* 2022;52(8):1560–8.
  46. Nouretdinov I, Costafreda SG, Gammerman A, Chervonenkis A, Vovk V, Vapnik V, et al. Machine learning classification with confidence: application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. *NeuroImage.* 2011;56(2):809–13.
  47. Sankar A, Zhang T, Gaonkar B, Doshi J, Erus G, Costafreda SG, et al. Diagnostic potential of structural neuroimaging for depression from a multi-ethnic community sample. *BJPsych Open.* 2016;2(4):247–54.
  48. Wise T, Marwood L, Perkins AM, Herane-Vives A, Joules R, Lythgoe DJ, et al. Instability of default mode network connectivity in major depression: a two-sample confirmation study. *Transl Psychiatry.* 2017;7(4):e1105.
  49. Wise T, Marwood L, Perkins AM, Herane-Vives A, Williams SCR, Young AH, et al. A morphometric signature of depressive symptoms in unmedicated patients with mood disorders. *Acta Psychiatr Scand.* 2018;138(1):73–82.
  50. Workman CI, Lythe KE, McKie S, Moll J, Gethin JA, Deakin JF, et al. Subgenual cingulate-amygdala functional disconnection and vulnerability to melancholic depression. *Neuropsychopharmacology.* 2016;41(8):2082–90.
  51. Ford BN, Yolken RH, Uppel RL, Teague TK, Irwin MR, Paulus MP, et al. Association of early-life stress with cytomegalovirus infection in adults with major depressive disorder. *JAMA Psychiatry.* 2019;76(5):545–7.
  52. Misaki M, Suzuki H, Savitz J, Drevets WC, Bodurka J. Individual variations in nucleus accumbens responses associated with major depressive disorder symptoms. *Sci Rep.* 2016;6(1):21227.
  53. Zheng H, Ford BN, Bergamino M, Kuplicki R, Hunt PW, Bodurka J, et al. A hidden menace? Cytomegalovirus infection is associated with reduced cortical gray matter volume in major depressive disorder. *Mol Psychiatry.* 2021;26(8):4234–44.
  54. Arnone D, McKie S, Elliott R, Thomas EJ, Downey D, Juhasz G, et al. Increased amygdala responses to sad but not fearful faces in major depression: relation to mood state and pharmacological treatment. *AJP.* 2012;169(8):841–50.
  55. Arnone D, McKie S, Elliott R, Juhasz G, Thomas EJ, Downey D, et al. State-dependent changes in hippocampal grey matter in depression. *Mol Psychiatry.* 2013;18(12):1265–72.
  56. Dutta A, McKie S, Downey D, Thomas E, Juhasz G, Arnone D, et al. Regional default mode network connectivity in major depressive disorder: modulation by acute intravenous citalopram. *Transl Psychiatry.* 2019;9(1):116.
  57. Godlewska BR, Browning M, Norbury R, Cowen PJ, Harmer CJ. Early changes in emotional processing as a marker of clinical response to SSRI treatment in depression. *Transl Psychiatry.* 2016;6(11):e957.
  58. Godlewska BR, Browning M, Norbury R, Igoumenou A, Cowen PJ, Harmer CJ. Predicting treatment response in depression: the role of anterior cingulate cortex. *Int J Neuropsychopharmacol.* 2018;21(11):988–96.
  59. Dunlop BW, Binder EB, Cubells JF, Goodman MM, Kelley ME, Kinkead B, et al. Predictors of remission in depression to individual and combined treatments (PRE-DICT): study protocol for a randomized controlled trial. *Trials.* 2012;13(1):106.
  60. Dunlop BW, Rajendra JK, Craighead WE, Kelley ME, McGrath CL, Choi KS, et al. Functional connectivity of the subcallosal cingulate cortex and differential outcomes to treatment with cognitive-behavioral therapy or antidepressant medication for major depressive disorder. *Am J Psychiatry.* 2017;174(6):533–45.
  61. Shi Y, Zhang L, Wang Z, Lu X, Wang T, Zhou D, et al. Multivariate machine learning analyses in identification of major depressive disorder using resting-state functional connectivity: a multicenter study. *ACS Chem Neurosci.* 2021;12(15):2878–86.
  62. Yan CG, Chen X, Li L, Castellanos FX, Bai TJ, Bo QJ, et al. Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proc Natl Acad Sci U S A.* 2019;116(18):9078–83.
  63. Hu J, Liu J, Liu Y, Wu X, Zhuang K, Chen Q, et al. Dysfunction of the anterior and intermediate hippocampal functional network in major depressive disorders across the adult lifespan. *Biol Psychol.* 2021;165:108192.
  64. Liu W, Wei D, Chen Q, Yang W, Meng J, Wu G, et al. Longitudinal test-retest neuroimaging data from healthy young adults in Southwest China. *Sci Data.* 2017;4(1):170017.
  65. Liu Y, Meng J, Wang K, Zhuang K, Chen Q, Yang W, et al. Morphometry of the hippocampus across the adult life-span in patients with depressive disorders: association with neuroticism. *Brain Topogr.* 2021;34(5):587–97.
  66. Sacchet MD, Livermore EE, Iglesias JE, Glover GH, Gotlib IH. Subcortical volumes differentiate major depressive disorder, bipolar disorder, and remitted major depressive disorder. *J Psychiatr Res.* 2015;68:91–8.
  67. Sacchet MD, Gotlib IH. Myelination of the brain in major depressive disorder: an in vivo quantitative magnetic resonance imaging study. *Sci Rep.* 2017;7(1):2200.
  68. Schwartz J, Ordaz SJ, Kircanski K, Ho TC, Davis EG, Camacho MC, et al. Resting-state functional connectivity and inflexibility of daily emotions in major depression. *J Affect Disord.* 2019;249:26–34.
  69. Navrady LB, Wolters MK, MacIntyre DJ, Clarke TK, Campbell AI, Murray AD, et al. Cohort profile: Stratifying resilience and depression longitudinally (STRADL): a questionnaire follow-up of generation Scotland: Scottish family health study (GS:SFHS). *Int J Epidemiol.* 2018;47(1):13–14g.
  70. Bomyea J, Choi SH, Sweet A, Stein M, Paulus M, Taylor C. Neural changes in reward processing following approach avoidance training for depression. *Soc Cogn Affect Neurosci.* 2022;17(3):336–49.
  71. Kryza-Lacombe M, Pearson N, Lyubomirsky S, Stein MB, Wiggins JL, Taylor CT. Changes in neural reward processing following amplification of positivity treatment for depression and anxiety: preliminary



- findings from a randomized waitlist controlled trial. *Behav Res Ther*. 2021;142:103860.
72. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56–62.
  73. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382–9.
  74. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003;54(5):573–83.
  75. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med*. 2004;34(1):73–82.
  76. Trajković G, Starčević V, Latas M, Leštarević M, Ille T, Bukumirić Z, et al. Reliability of the Hamilton rating scale for depression: a meta-analysis over a period of 49 years. *Psychiatry Res*. 2011;189(1):1–9.
  77. Leucht S, Fennema H, Engel RR, Kaspers-Janssen M, Szegedi A. Translating the HAM-D into the MADRS and vice versa with equipercentile linking. *J Affect Disord*. 2018;226:326–31.
  78. Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, et al. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. *Neuropsychopharmacol*. 2005;30(2):405–16.
  79. Uher R, Farmer A, Maier W, Rietschel M, Hauser J, Marusic A, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med*. 2008;38(2):289–300.
  80. Bernstein IH, Rush AJ, Stegman D, Macleod L, Witte B, Trivedi MH. A comparison of the QIDS-C<sub>16</sub>, QIDS-SR<sub>16</sub>, and the MADRS in an adult outpatient clinical sample. *CNS Spectr*. 2010;15(7):458–68.
  81. Liu J, Xiang YT, Lei H, Wang Q, Wang G, Ungvari GS, et al. Guidance on the conversion of the Chinese versions of the Quick Inventory of Depressive Symptomatology-Self-Report (C-QIDS-SR) and the Montgomery-Asberg Scale (C-MADRS) in Chinese patients with major depression. *J Affect Disord*. 2014;152–154:530–3.
  82. Citrome L, Abi-Dargham A, Bilder RM, Duffy RA, Dunlop BW, Harvey PD, et al. Making sense of the matrix: a qualitative assessment and commentary on connecting psychiatric symptom scale items to the research domain criteria (RDoC). *Innov Clin Neurosci*. 2022;19(1–3):26–32.
  83. Cuthbert BN. Research domain criteria: toward future psychiatric nosologies. *Dialogues Clin Neurosci*. 2015;17(1):89–97.
  84. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *AJP*. 2010;167(7):748–51.
  85. Uher R, Frey BN, Quilty LC, Rotzinger S, Blier P, Foster JA, et al. Symptom dimension of interest-activity indicates need for aripiprazole augmentation of escitalopram in major depressive disorder: a CAN-BIND-1 report. *J Clin Psychiatry*. 2020;81(4):9256.
  86. Furukawa TA, Streiner DL, Azuma H, Higuchi T, Kamijima K, Kanba S, et al. Cross-cultural equivalence in depression assessment: Japan-Europe-north American study. *Acta Psychiatr Scand*. 2005;112(4):279–85.
  87. Vindbjerg E, Makransky G, Mortensen EL, Carlsson J. Cross-cultural psychometric properties of the Hamilton depression rating scale. *Can J Psychiatr*. 2019;64(1):39–46.
  88. Bashyam VM, Doshi J, Erus G, Srinivasan D, Abdulkadir A, Singh A, et al. Deep generative medical image harmonization for improving cross-site generalization in deep learning predictors. *J Magn Reson Imaging*. 2021;55(3):908–16.
  89. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010;29(6):1310–20.
  90. Doshi J, Erus G, Ou Y, Gaonkar B, Davatzikos C. Multi-atlas skull-stripping. *Acad Radiol*. 2013;20(12):1566–76.
  91. Doshi J, Erus G, Ou Y, Resnick SM, Gur RC, Gur RE, et al. MUSE: Multi-atlas region segmentation utilizing ensembles of registration algorithms and parameters, and locally optimal atlas selection. *NeuroImage*. 2016;127:186–95.
  92. Asman AJ, Chambless LB, Thompson RC, Landman BA. Out-of-atlas likelihood estimation using multi-atlas segmentation. *Med Phys*. 2013;40(4):043702.
  93. Davatzikos C, Genc A, Xu D, Resnick SM. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage*. 2001;14(6):1361–9.
  94. Ou Y, Sotiras A, Paragios N, Davatzikos C. DRAMMS: deformable registration via attribute matching and mutual-saliency weighting. *Med Image Anal*. 2011;15(4):622–39.
  95. Sotiras A, Resnick SM, Davatzikos C. Finding imaging patterns of structural covariance via non-negative matrix factorization. *NeuroImage*. 2015;108:1–16.
  96. Sotiras A, Toledo JB, Gur RE, Gur RC, Satterthwaite TD, Davatzikos C. Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. *Proc Natl Acad Sci U S A*. 2017;114(13):3527–32.
  97. Wen J, Varol E, Chand G, Sotiras A, Davatzikos C. MAGIC: Multi-scale heterogeneity analysis and clustering for brain diseases. In: Martel AL, Abolmaesumi P, Stoyanov D, Mateus D, Zuluaga MA, Zhou SK, et al., editors. Medical image computing and computer assisted intervention – MICCAI 2020. Cham: Springer International Publishing; 2020. p. 678–87. (Lecture Notes in Computer Science).
  98. Li H, Satterthwaite TD, Fan Y. Large-scale sparse functional networks from resting state fMRI. *NeuroImage*. 2017;156:1–13.
  99. Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, et al. Prediction of individual brain maturity using fMRI. *Science*. 2010;329(5997):1358–61.
  100. Jafri MJ, Pearson GD, Stevens M, Calhoun VD. A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*. 2008;39(4):1666–81.
  101. Savadjiev P, Kindlmann GL, Bouix S, Shenton ME, Westin CF. Local white matter geometry from diffusion tensor gradients. *NeuroImage*. 2010;49(4):3175–86.
  102. Zeng LL, Shen H, Liu L, Wang L, Li B, Fang P, et al. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain*. 2012;135(5):1498–507.
  103. Toloşi L, Lengauer T. Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics*. 2011;27(14):1986–94.
  104. Elad M. Sparse and redundant representations: from theory to applications in signal and image processing. New York: Springer; 2010. p. 376.
  105. Cui Z, Li H, Xia CH, Larsen B, Adebimpe A, Baum GL, et al. Individual variation in functional topography of association networks in youth. *Neuron*. 2020 Apr;106(2):340–353.e8.
  106. Kretz-Delgado K, Murray JF, Rao BD, Engan K, Lee TW, Sejnowski TJ. Dictionary learning algorithms for sparse representation. *Neural Comput*. 2003;15(2):349–96.
  107. Olshausen BA, Field DJ. Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vis Res*. 1997;37(23):3311–25.
  108. Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*. 2000;287(5456):1273–6.
  109. Eavani H, Satterthwaite TD, Filipovych R, Gur RE, Gur RC, Davatzikos C. Identifying sparse connectivity patterns in the brain using resting-state fMRI. *NeuroImage*. 2015;105:286–99.
  110. Sahoo D, Satterthwaite TD, Davatzikos C. Hierarchical extraction of functional connectivity components in human brain using resting-state fMRI. *IEEE Trans Med Imaging*. 2021;40(3):940–50.
  111. Beer JC, Tustison NJ, Cook PA, Davatzikos C, Sheline YI, Shinohara RT, et al. Longitudinal ComBat: a method for harmonizing longitudinal multi-scanner imaging data. *NeuroImage*. 2020;220:117129.
  112. Fortin JP, Sweeney EM, Muschelli J, Crainiceanu CM, Shinohara RT. Alzheimer's Disease Neuroimaging Initiative. Removing inter-subject technical variability in magnetic resonance imaging studies. *Neuroimage*. 2016;132:198–212.
  113. Fortin JP, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*. 2018;167:104–20.
  114. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8(1):118–27.

115. Chen AA, Srinivasan D, Pomponio R, Fan Y, Nasrallah IM, Resnick SM, et al. Harmonizing functional connectivity reduces scanner effects in community detection. *NeuroImage*. 2022;256:119198.
116. Chen AA, Beer JC, Tustison NJ, Cook PA, Shinohara RT, Shou H, et al. Mitigating site effects in covariance for machine learning in neuroimaging data. *Hum Brain Mapp*. 2022;43(4):1179–95.
117. McDaid AF, Joshi PK, Porcu E, Komljenovic A, Li H, Sorrentino V, et al. Bayesian association scan reveals loci associated with human lifespan and linked biomarkers. *Nat Commun*. 2017;8:15842.
118. Wen J, Varol E, Sotiras A, Yang Z, Chand GB, Erus G, et al. Multi-scale semi-supervised clustering of brain images: deriving disease subtypes. *Med Image Anal*. 2022;1(75):102304.
119. Yang Z, Wen J, Davatzikos C. Smile-GANs: Semi-supervised clustering via GANs for dissecting brain disease heterogeneity from medical images. arXiv:2006.15255 [cs, eess, q-bio, stat]. 2020; Cited 2022 Apr 29. Available from: <http://arxiv.org/abs/2006.15255>.
120. Chand GB, Dwyer DB, Erus G, Sotiras A, Varol E, Srinivasan D, et al. Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain*. 2020;143(3):1027–38.
121. Timmerby N, Andersen JH, Søndergaard S, Østergaard SD, Bech P. A systematic review of the clinimetric properties of the 6-item version of the Hamilton depression rating scale (HAM-D6). *Psychother Psychosom*. 2017;86(3):141–9.

### Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

