Multi-Modal Abnormalities of Brain Structure and Function in Major Depressive Disorder: A Meta-Analysis of Neuroimaging Studies

Jodie P. Gray, MS; Veronika I. Müller, PhD; Simon B. Eickhoff, MD; Peter T. Fox, MD

Previous presentation: None.

Corresponding authors and address:

Jodie Gray, MS

Peter T. Fox, MD

Research Imaging Institute, University of Texas Health Science Center at San Antonio

7703 Floyd Curl Drive, San Antonio, TX 78229

Tel: (210) 567-8100

Email: jodiepgray@gmail.com, fox@uthscsa.edu

Disclosures: Ms. Gray has received research-training support from a grant through the US Department of Defense; she has received honoraria for speaking from the Mind Science Foundation. Dr. Eickhoff has received research support from the Deutsche Forschungsgemeinschaft, The Helmholtz-Association, The National Institutes of Health, and the European Commission Dr. Fox has received research support from The National Institutes of Health, the US Department of Defense, the US Department of Veterans Affairs, the Cancer Prevention & Research Institute of Texas, and The Mather's Foundation. Dr. Müller reports no financial relationships with commercial interests.

Acknowledgements: This work was supported by grants from the National Institute of Mental Health (R01-MH074457-14) and the Department of Defense (ISG/W81XWH1320065).

Abstract

Objective: Imaging studies of major depressive disorder (MDD) have reported structural and functional abnormalities in many, spatially diverse brain regions. Quantitative meta-analyses of this literature, however, have failed to find statistically significant between-study spatial convergence, other than transdiagnostic-only effects. In the present study, the authors apply a novel, multi-modal, meta-analytic approach to test the hypothesis that MDD exhibits spatially convergent structural and functional brain abnormalities.

Methods: This coordinate-based meta-analysis (CBMA) included voxel-based morphometry (VBM) studies and resting-state voxel-based pathophysiology (VBP) studies imaging blood flow (BF), glucose metabolism, regional homogeneity (ReHo), and amplitude of low frequency fluctuations (ALFF/fALFF). Input data were grouped into three primary meta-analytic classes: gray matter atrophy; increased function; and, decreased function in MDD patients relative to healthy controls. Secondary meta-analyses grouped across primary categories. Tertiary analyses grouped by medication status and absence of psychiatric comorbidity. Activation likelihood estimation (ALE) was used for all analyses.

Results: In total 92 publications reporting 152 experiments were identified, collectively representing 2,928 MDD patients. Primary analyses detected no convergence across studies. Secondary analyses identified portions of subgenual cingulate, hippocampus, amygdala, putamen, retrosplenial cortex, and middle occipital/inferior temporal gyri as demonstrating convergent abnormalities. Tertiary analyses (clinical subtypes) showed improved convergence relative to secondary analyses.

Conclusions: CBMA identified spatially convergent structural (VBM) and functional (VBP) abnormalities in MDD. Present findings suggest replicable neuroimaging features associated with MDD, beyond the transdiagnostic effects reported in prior meta-analysis.

Our findings support continued research focus on the subgenual cingulate and other select regions' role in MDD.

Introduction

Major depressive disorder (MDD) is the single largest contributor to disability worldwide, impacting as many as 300 million sufferers annually (1). Despite decades of basic science, clinical neuroscience, and psychiatric research, the pathophysiology of MDD is not well understood (2). Human neuroimaging approaches comprise powerful, non-invasive methods to investigate the neurobiological mechanisms underpinning psychiatric disorders (3,4). Neuroimaging's promise notwithstanding, recent reports have challenged the reliability of this literature, drawing attention to small sample sizes (5), clinical heterogeneity (6), and flawed correction for multiple comparisons (7), which jointly work to inflate false positive rates. Though improvement of neuroimaging techniques is currently an active area of research, previous findings are not without value. Meta-analytic approaches are capable of addressing many of the methodological concerns that contribute to varied findings at the individual study level and allow identification of reliable, true-positive findings in existing literature.

Coordinate-based meta-analysis (CBMA) is a well-established family of methods which hold a prominent position in neuroimaging research (8, 9). CBMA offers a large-scale, data-driven approach to the identification of brain regions consistently altered by disease by testing for spatial convergence across reported findings from previously published neuroimaging studies. CBMA tests for convergence against the null hypothesis that reported findings follow a random spatial distribution across the brain, rather than demonstrating convergent abnormality in discrete brain regions. CBMA is applicable only to data acquired from the whole brain and analyzed in voxel-wise manner, to ensure identification of convergent effects in a spatially unbiased (non-ROI based) manner (8, 10). CBMA applies equally well to multiple types of imaging data including task activation, voxel-based morphometry (VBM; 11), and resting state voxel-

based pathophysiology (VBP). Activation/anatomical likelihood estimation (ALE; 10, 12-14) — the most widely adopted CBMA method (15) — computes the union of reported findings based entirely on location. Unlike effect-size meta-analysis, ALE is blind to magnitude and sign (+/-) of effect. Though ALE has been traditionally employed in single modality meta-analysis, its flexibility for integration of findings across imaging methods allows for comprehensive assessment of disease related effects. In the present study, we employed CBMA/ALE to identify convergent structural (VBM) and physiological (VBP) disease effects of both signs (+/-).

Several recent, large-scale meta-analyses of structural and functional imaging in MDD and other psychiatric disorders strongly suggest that concerns about a preponderance of false positives in the neuroimaging literature are well justified. Table 1 provides an outline of the meta-analyses described in the present summary. A transdiagnostic meta-analysis of task-activation studies across several Axis-I disorders by Sprooten and colleagues (16) found shared effects across diseases, however no effect of diagnosis or RDoC domain on spatial distribution of reported findings. A similar transdiagnostic ALE meta-analysis by Goodkind and colleagues (17) assessed regional atrophy (VBM) patterns across 6 Axis-I disorders and did find spatial convergences in bilateral anterior insula and anterior cingulate when assessing across all disorders, but found no unique characteristics of any disorder. A transdiagnostic independent component analysis (ICA) meta-analysis of VBM studies Vanasse and colleagues (our laboratory) assessed disease loadings on several independent brain networks (18). The central finding from Vanasse et al was that no disease loaded on a single network, and no network loaded on a single disease. Furthermore, one of the component networks identified by Vanasse's group closely reflected the pattern of shared pathology identified through Goodkind's transdiagnostic ALE analysis of VBM data. An ALE meta-analysis of resting-state VBP studies across 11 neuropsychiatric disorders by Sha and colleagues

(19) also identified shared effects across diseases, including widespread abnormalities in MDD patients. However, the disease-specific distribution identified in Sha et al.'s analysis of MDD failed to converge at statistical thresholds recommended by ALE best-practice guidelines (20, 21). In an ALE meta-analysis of cognitive and emotional task-activation studies limited to MDD cohorts, Müller et al (our group; 6), no brain regions of significant convergence were identified. Collectively, these meta-analyses pose a challenge to the psychiatric neuroimaging community.

Findings from the aforementioned transdiagnostic studies strongly indicate shared pathology across neuropsychiatric diseases and weak neurobiological "signal" of depression alone – the primary finding of the MDD-specific meta-analysis by Müller et al. Shared pathology across illnesses is "not a component of current psychiatric nosology" (Goodkind et al, 16), though these findings may align with newer research initiatives such as the Research Domain Criteria (RDoC) project (22, 23). Despite the notable successes of transdiagnostic meta-analyses and absence of significant single-diagnosis findings, disease-specific approaches remain an important area for research. The mental health care system broadly considered —care providers, insurance providers, regulatory agencies— are not likely to abandon established psychiatric terminology in light of the above neurobiological observations. Further, clinical trials testing new therapies are typically carried out in patients falling into specific diagnostic categories, rather than being symptom-driven or transdiagnostic. For these reasons, it is important to determine if MDD-specific, regional neurobiological changes can be detected using neuroimaging. Prior studies have failed to find a neurobiological "signature" of MDD alone using task-activation only (6,16), VBM only (17,18), and resting-state VBP only (19). We propose to test this hypothesis using both VBM and resting-state VBP studies in combination.

The co-localization of structural and functional and structural abnormalities is well documented across neuropsychiatric diseases. Numerous disorders including

Parkinson's disease (24), Alzheimer's disease (25), primary progressive apraxia (26), multiple sclerosis (27, 28), schizophrenia (29), and various mood disorders including MDD and bipolar disorder (30, 31), have demonstrated conjoint abnormalities of brain function and structure, with recent research investigating this relationship in MDD specifically (32, 33). The concordance of structural and functional abnormalities in neurodegenerative diseases underlies the network degeneration hypothesis (34-36), of which MDD is also being investigated as a potentially network-based disorder (37-40). Furthermore, recent research investigating the network degeneration hypothesis indicates that high traffic network "hubs" are more likely to experience gray matter lesions as a result of disease-related overstimulation (41, 42), which may contribute to subsequent decreases in brain function at affected regions. As we anticipate the co-localization of structural and functional disease effects in MDD, a central objective of the present study is to examine the convergence of gray matter atrophy, increases, and decreases in resting-state function both independently and jointly. Currently, voxel-based morphometry (VBM) investigations of gray matter alterations in MDD form a large body of literature suitable for meta-analysis. Similarly, voxel-based physiological (VBP) investigations using PET and SPECT imaging of brain blood flow and glucose metabolism, together with recent advances in functional magnetic resonance imaging techniques contribute to a growing corpus of VBP literature in MDD. The present study is among the first to comprehensively assess resting-state functional (VBP) and structural (VBM) findings in MDD both independently and in pooled multi-modal datasets.

Thus, the objective of this meta-analysis is to assess the spatial convergence of brain abnormalities in MDD as detected by structural and resting-state functional neuroimaging data. Our primary hypothesis is that MDD will demonstrate pathological changes detectable across neuroimaging paradigms. We hypothesize localized convergence of: (a) gray matter atrophy, (b) increased, and (c) decreased brain function

in MDD patients relative to controls. We also hypothesize improved co-localization of abnormalities, as evaluated through pooled datasets for secondary analysis (d). The meta-analytic design and statistical thresholds for the present study were selected to emulate Müller et al's 2017 study in order to compare the findings from task-based versus task-independent investigations of MDD. We also hypothesize that accounting for the clinical heterogeneity of MDD, by assembling patient subgroups (to the degree possible through available literature), can enhance the convergence of identified brain regions. Hypotheses confirmed by this meta-analytic approach, we would submit, should be regarded as providing direction for further primary-data studies, rather than seen as established conclusions.

Methods

Literature Search

A literature search of PubMed, Google Scholar, BrainMap (18, 43-45) and reference tracing of previous meta-analyses was performed to identify MDD neuroimaging experiments reporting either gray matter atrophy, or increased resting state function, or decreased resting state function compared to healthy control subjects. MDD related hypertrophy, a rare phenomenon occasionally reported in remitted MDD (relative to acute MDD), was not included in this analysis. Voxel-based morphometry (VBM) studies and resting state VPB studies of regional cerebral blood flow (rCBF), regional homogeneity (ReHo), amplitude of low frequency fluctuations (ALFF/fALFF), and regional glucose metabolism were identified using various combinations of the search terms *major depressive disorder, major depression, depression, unipolar depression, VBM, gray matter, rCBF, positron emission tomography (PET), single photon emission computed tomography (SPECT), arterial spin labeling (ASL), ReHo, ALFF/fALFF, glucose metabolism, brain activity, and resting state.* The literature search was

completed January 2018. A study selection diagram for this meta-analysis is detailed in Figure 1 and further details of the literature search are provided in Supplemental Appendix 1.

Study selection criteria pertaining to indices of quality

Preliminary selection criteria required that studies be peer-reviewed, English language neuroimaging reports. Studies identified outside of the BrainMap database were reviewed by BrainMap team members and subsequently coded through Scribe and submitted to the database (21). Standard expectations for publication in this field is for application of motion correction, measures to limit motion during scan, and/or exclusion of data that exhibited excessive motion during acquisition. Measure(s) for motion correction utilized in each included study are tabulated in ST1 the Supplemental Materials.

Study selection criteria relating to subjects

MDD patients from included studies were diagnosed using DSM-III (4 studies), DSM-IV (85 studies), or ICD-10 (3 studies) evaluation. Only studies comparing patients in the acute phase of MDD to healthy controls were included. Experiments including remitted subjects (n=2) or any contrast other than MDD vs. healthy controls were excluded (n=5)(Figure 1). Studies utilizing dual diagnosis patient populations with other major medical illness (e.g., MDD and hypothyroidism) or psychiatric comorbidities were excluded (n=13). However, we allowed for the inclusion of studies in which partial populations of the patient cohort had comorbidities (e.g., subset of MDD patients with anxiety symptoms) with the criterion that MDD was the primary diagnosis. Studies with strict exclusion criteria for psychiatric comorbidities were flagged for use in subsequent meta-analytic grouping. We allowed for the inclusion of studies that featured patient populations of varying medication status, but flagged those studies that recruited patient

populations of specific medication status (all medicated, treatment naive, drug washout) for subsequent meta-analytic grouping. We also flagged studies that recruited patient populations of specific severity or disease onset (first episode, chronic/recurrent, treatment resistant, adolescent, geriatric).

Study selection criteria relating to technical aspects

Studies of resting-state VBP included investigations of regional cerebral blood flow (rCBF), regional cerebral glucose metabolism, regional homogeneity (ReHo), and amplitude of low frequency fluctuations (ALFF/fALFF) using imaging methods of positron emission tomography (PET), single photon emission computed tomography (SPECT), and functional magnetic resonance imaging (fMRI) (Figure 1). Included resting-state VBP studies only allowed for those that used voxel-wise whole-brain methods to compare MDD patients to healthy control subjects. Thirty-six studies investigating functional or effective connectivity were excluded from this meta-analysis because they used regional sampling (N=26), used incompatible patient-group contrasts (N=2), were review articles or meta-analyses (N=4), or were multivariate analyses only, without mass-univariate analyses (N=2), and the remaining studies (N=2) were excluded because as functional or effective connectivity studies they cannot be integrated in current coordinate-based meta-analysis methods.

Included studies of gray matter volume utilized voxel-based morphometry methods. Included studies only allowed for those that used voxel-wise whole-brain methods to compare MDD patients to healthy control subjects. Studies using non whole-brain methods, such as ROI or network restricted sampling, (n=19) were excluded.

Only studies reporting results as coordinates using standard reference space (Talairach or Montreal Neurological Institute [MNI]) were included; those studies which did not report results in the form of standardized coordinates (n=8) or did not report

coordinate system used (n=1) were excluded. Coordinates were converted to Talairach space for this analysis (46). To avoid repeated inclusion of the same patient populations we carefully screened studies that pulled from open source or national data repositories and excluded those that reported use of a patient cohort already included in this meta-analysis (n=4).

Data non-redundancy is of crucial importance to avoid bias in meta-analytic findings. To avoid inclusion of duplicate patient populations in the present study, we performed several pre- and post hoc assessments. First, we carefully screened studies that pulled from open source or national data repositories and excluded those that reported duplicate patient cohorts (n=4). For multiple studies deriving from the same research group, patient populations and reported coordinates were inspected to assess potential redundancy (18). For studies reporting multiple contrasts from the same patient population, only one contrast per patient population was used in each meta-analysis (47). Post-hoc assessments were performed in cases when multiple experiments from the same research group contributed to identified clusters. In these cases, leave-one-out analyses were performed to assess potential redundant contributions to identified clusters. In cases when potential patient overlap was indicated by leave-one-out analysis, only the largest study was included for final meta-analysis. Another meta-analysis best practice is to contact authors in confirm to data independence, however, attempts to contact research groups of interest are not always successful.

All Effects Analysis

Coordinates from all included studies were collectively pooled to generate a unified **All Effects** meta-analytic category. Coordinates from multiple experimental contrasts obtained from the same subject group (such as studies that reported both locations of gray matter atrophy and locations of increased or decreased function relative to controls

in separate experiments) were concatenated to generate sign-independent foci groups for each patient population tested. For detailed description, see SA1 for “Details of ALE Analysis in Supplemental Materials. All patients, including both medicated and unmedicated patients at time of scanning, were included in the All Effects analysis.

Meta-analytic Data Classes

To assess modality-specific contributions to findings from the All Effects analysis, input data were grouped into primary and secondary meta-analytic groups for analysis. Primary analyses included three single modality classes and two dual modality classes, as follows. Three single modality classes were created by grouping experiments of decreased gray matter volume in MDD patients compared to controls (**VBMneg**), decrease in resting state function in MDD patients compared to controls (**VBPneg**), and increased function in MDD patients compared to control subjects (**VBPpos**) (Figure 2- items 1, 2, & 3 respectively); two dual modality classes were created by combining the classes of rsFX-decrease and rsFX-increase with the GMV data (**VBPneg + VBMneg** and **VBPpos + VBMneg**) (Figure 2- items 4 & 5, respectively). In studies that reported both VBM and VBP changes in the same subject population, we included only the coordinates reporting change in VBP in the pooled datasets (n=5). For initial analysis of the five major meta-analytic classes, all available data (including patients of varying medication status at time of scan) were included to test convergence of clinically heterogeneous patient groups.

Patient Groupings

Studies were grouped for analysis into two tiers: All Effects (all patient types) and patient subgroups. Studies were subgrouped as: a) drug/treatment naive MDD only; b) treated MDD with drug washout before imaging, and c) MDD with no psychiatric comorbidities.

For each sub-group, all five classes of experiments were analyzed, provided the number of included experiments was sufficient for robust ALE calculation (Figure 2- items 6 through 18). Per Eickhoff et al 2016, the minimum number of experiments required for robust ALE analysis is 17 (20).

Other subgroups attempted (first episode, chronic/recurrent, treatment resistant, adolescent, geriatric) did not include a sufficient number of individual experiments for meta-analysis (see Discussion). Drug washout groups also did not include a sufficient number of experiments for standalone analysis. In an effort to maximize the use of available information from individual studies, we combined the drug washout and drug naive groups to assess potential effects from a medication-free group.

ALE meta-analysis

Activation likelihood estimation (ALE) (8,10,14,20) was performed using GingerALE (version 3.0) software (48). The ALE algorithm was originally developed for use in task-activation functional studies (12), but has undergone numerous revisions including adaptations for use with VBM studies (47,49). ALE assesses spatial convergence of reported findings against the null hypothesis that findings follow a random spatial distribution rather than demonstrating statistically significant convergence at discrete regions. The most current versions of ALE model reported coordinates, or foci, as 3D Gaussian probability distributions to generate per-experiment modeled-activation (or modeled-atrophy) maps (50). ALE derives full-width half-maximum for each Gaussian distribution based on sample size, allowing experiments with larger subject sizes greater statistical certainty. ALE generates a union map of all per-experiment MA maps and tests for above-chance spatial convergence through a variety of available thresholding options. A revised version of the algorithm (48) recommends either Family-wise error or Cluster-level inference thresholding methods for robust analysis. The selected method

for the present study, cluster-level inference, generates a simulated dataset of randomly distributed foci based on characteristics of the input dataset for testing the null hypothesis.

Results were thresholded for significance using cluster-level inference of $P<0.05$ with a cluster-forming threshold of $P<0.001$ to reflect the study design from Müller's 2017 study (6) and ALE best-practices (21, 51). ALE analysis was re-tested at cluster-level inference of $P<0.0027$ (Bonferroni correction for multiple comparisons of 18 total meta-analyses) to assess the robustness of identified clusters.

Noise simulation for estimation of file-drawer effect

Presently, the potential for unpublished null findings in the neuroimaging literature is not accounted for in the ALE algorithm, as ALE's focus is to assess convergence of non-null findings of which a large portion are anticipated to be false positives (52). Potential publication bias in the present study was evaluated through a modified version of the Fail-Safe N method by Acar et al to estimate the robustness of identified results against unpublished neuroimaging findings (15). A recent simulation utilizing the BrainMap database identified that missing contrasts may be estimated at 6 per every 100 instances of reported findings (53). Thus, we re-tested convergent meta-analyses with an additional 6% added noise to assess the robustness of identified clusters. Surviving clusters were subsequently re-tested with higher rates of noise up to 30%.

Results

A total of 92 papers (97 studies) with 152 individual experiments comprising results from 2,928 patients were identified for inclusion in this meta-analysis. The number of experiments included in each major meta-analytic category were: VBMneg: 43 experiments; VBPneg: 62 experiments; and, VBPpos: 47 experiments. ST1 & ST2 list all

studies included in meta-analysis. SF1 shows the distribution of foci from each of the major meta-analytic categories (see Supplemental Materials).

All Effects Analysis

The All Effects analysis comprised a total of 102 foci groups created from summed results from all experiment types. This unified analysis identified a single region within the left hippocampus as demonstrating convergent abnormality, listed in Table 2 and shown in Figure 3.

Heterogeneous Group: All Patients

Among the classes of GMV, VBPneg, VBPpos, and VBPneg + VBMneg utilizing all pooled patient data, none revealed any significant regions of convergent brain abnormality. The class of VBPpos + VBMneg utilizing clinically heterogeneous patient data identified consistent aberrant brain regions in MDD within the left hippocampus (as identified in the All Effects analysis) and an additional region of significant convergence in the subgenual cingulate cortex (listed in Table 2 and illustrated in Figure 3).

Patient Subgroups

Clinical subgroups that fulfilled n>17 experiments criteria included the categories of drug naive patients (**Drug Naive)**, drug naive and drug washout patients combined (**Naive+Washout**), and those studies with strict exclusion criteria of comorbid psychiatric disorders for MDD patients (**Only MDD**). See ST3-8 in supplement for lists of studies included in each meta-analytic grouping.

A total of 13 clinical subgroups across the five meta-analytic classes included a sufficient number of experiments to perform ALE analysis. Among the 13 subgroups, only 5 yielded significant results (listed in Table 2 and illustrated in Figure 3). Clinical subgroups within the VBPpos + VBMneg class identified significant convergence among

the Drug Naive, Naive+Washout, and the Only MDD clinical groups. Consistent abnormal brain areas identified within these clinical subgroups included regions of the left hippocampus (as previously identified) and an additional region including areas of the right amygdala and ventral anterior putamen. Clinical subgroups from the VBPneg class identified significant regions among the Drug Naive and Only MDD clinical groups. Areas identified within these clinical subgroups included three additional regions to those previously identified, a region encompassing areas of the left middle occipital and left inferior temporal gyri, a region within the left retrosplenial cortex, and a region within the right putamen.

Convergence by Imaging Modality

Data from various imaging modalities contributed to the clusters identified through ALE analysis, with no single modality being profoundly over-represented in any result from the VBPpos + VBMneg class. Regions identified in clinical subgroups of the VBPneg class were somewhat dominated by contributions from ALFF and ReHo experiments, though this could be attributed to their overall representation within the dataset tested. For detailed distributions, see ST9-11.

Noise Simulation

Each identified cluster was re-tested in meta-analyses with added noise (beginning at 6% noise) to assess robustness against potentially unpublished findings. Surviving clusters were subsequently re-tested with higher rates of noise up to 30%. Table 2 details the Fail Safe N percentage (FSN %) of additional noise that must be added to each meta-analysis to result in failure of convergence for previously identified clusters. In general, noisier contrasts (e.g., All Effect and other groups including All Patients) are less robust against the simulation of additional noise.

Discussion

Our findings do not support our first three hypotheses, with no brain regions of significant convergence arising from meta-analysis of (a) gray matter atrophy, (b) increased, or (c) decreased brain function in MDD patients compared to controls. However, brain regions demonstrating significant abnormality did arise from meta-analysis of pooled structure-function findings in MDD, supporting our hypothesis of co-localized effects (d). We also identified additional regions of convergence from meta-analysis of clinical subgroups despite decreased sample size.

The present work, to our knowledge, is the first to comprehensively assess multi-modal imaging data to investigate the convergence of voxel-based morphometry and voxel-based pathophysiology findings in MDD. Regions of significant convergence identified in this study include: the subgenual cingulate cortex, the left hippocampus, the right amygdala/putamen, the left retrosplenial cortex, and the right middle occipital/inferior temporal gyri. The brain regions identified in the present meta-analysis are included in many current models of MDD pathology and treatment approaches. Furthermore, our methods of largely pooling multi-modal data, and conversely delineating data by available clinical details, improved convergence of results. Our identification of brain regions demonstrating reliable abnormality in MDD, whereas previous meta-analyses have failed to identify any disease-specific effects in MDD, is a significant contribution to existing literature. We view these results as motivations for refinement of future primary studies in MDD.

Identified Regions

Identification of consistent abnormality within the subgenual anterior cingulate cortex (hereafter subgenual cingulate) in the present meta-analysis is a potentially important finding for the current state of MDD research. The subgenual cingulate has been widely

implicated in major depressive disorder as a regulator of mood (30, 31, 54-56), in the processing of emotional stimuli (57-59) and as a target for network based treatments such as deep brain stimulation (60- 63) and a downstream target for transcranial magnetic stimulation (64-67). Reliable identification of the subgenual cingulate through large-scale, multi-modal meta-analysis strongly supports further research of this region's role in MDD. Furthermore, the subgenual cingulate has not been reliably identified in transdiagnostic meta-analyses such as the VBM-ALE analysis conducted by Goodkind et al (17) (although subgenual cingulate does appear to be present in the all-groups analysis; see Figure 2a). This distinctive finding arising from the present study suggests that disease-specific effects, beyond transdiagnostic-only effects, are detectable in neuroimaging data and warrants further exploration.

The left hippocampus was also identified in this study. Decreased hippocampal volume has been observed in neuroimaging studies of MDD over the past 20 years (68-72). Hypotheses of MDD-related hippocampal volume decline posit that the hippocampus may be affected by stress (72) and may contribute to the cognitive (73) and recollection memory deficiencies (74, 75) often present in MDD sufferers. The hippocampus has also been implicated in MDD through disrupted hippocampal connectivity effects on self-referential activity in MDD (76), and demonstrated conjoint reductions in gray matter density and activation during working memory task in MDD patients (77). The identification of the hippocampus in the present study is a notable finding for both past and future investigations of this brain region's role in MDD.

We also identified regions of the right amygdala and right putamen in the present study. In recent investigations utilizing emotional valence paradigms, the amygdala has demonstrated aberrant activation in MDD patients compared to healthy controls (78). The amygdala has also demonstrated reliable volume differences in unmedicated MDD patients relative to controls in a meta-analysis of 13 individual neuroimaging studies

(79). The putamen, though its potential role is less well established in MDD, has also demonstrated volumetric and shape abnormalities in untreated, first episode MDD (80). The putamen has also demonstrated functional disruption in MDD through investigation of the correlation between anhedonia severity and aberrant neural activity in response to emotional stimuli (81). Our findings suggest the need for continued investigation of the amygdala and putamen's potential roles in MDD.

Other regions identified in this study including the left retrosplenial cortex (encompassing regions within both BA 29 and BA 30), and an overlapping area of the right middle occipital and inferior temporal gyri currently do not have well established roles in current models of MDD. A recent, large study of MDD patients (n=336) demonstrated that altered functional connectivity between the retrosplenial cortex and other key brain regions may contribute to increased rumination symptoms in depression (82). Brain perfusion deficits in occipital areas have been previously observed in adolescent MDD patients, however, this study provided uncertain conclusions for the region's significance in MDD (83). More recently, increased functional connectivity with the right middle occipital gyrus and the amygdala has been observed in association with cognitive dysfunction in MDD (84). The right middle occipital/ inferior temporal gyri have also demonstrated reduced cortical thickness in MDD patients compared to controls in a recent large-scale study from the ENIGMA cohort (N=1902) (85). Though the role of these regions in MDD is less well defined, results from the present study indicate that further investigation of these regions' potential role in the pathophysiology of MDD is warranted.

Convergence from Patient Groupings

Separation of data into patient subgroups played a critical role in identifying additional brain regions beyond those found in the more heterogeneous groups. To our knowledge,

this is the largest meta-analysis in MDD that also included a sufficient number of experiments to perform ALE analysis in subgroups. Our findings suggest that clinical heterogeneity in MDD has observable neuroimaging effects, which warrant further investigation.

Our study was limited by the recruiting and reporting methods employed at the individual study level. Author-defined patient groups in the present study were largely limited to medication status (treatment naive and drug washout groups). Other categories including patient groups of specific severity (first episode, recurrent/chronic, treatment resistant) and age of onset (geriatric or adolescent depression) did not yield n>17 experiments (20) and were not further analyzed in our study. Of note, 60% of the experiments included in the Drug Naive categories (VBPpos + VBMneg and VBPneg Drug Naïve groups) included experiments reporting findings from first-episode MDD patients (see ST5 and ST8). Findings from these subgroups may indicate neuroimaging effects specific to first-episode MDD, though we were not able to reliably test this effect in the present analysis. As such, we strongly recommend future studies investigating this potential effect.

More meaningful patient categories for future studies would ideally focus on severity, duration, and treatment response in MDD. Of the 92 publications included in the present analysis, 42 studies (51% of total) recruited mixed MDD patient populations and pooled all patients into heterogeneous groups regardless of age of onset, disease duration/severity, and number of previous episodes. It is a common convention in neuroimaging publications to include patient demographic tables reporting the mean and standard deviation of the aforementioned clinical features, though this is not standardized or consistent among current literature. A central recommendation from the present work is for standardized recruiting and reporting mechanisms to be adopted at the individual study level. Due to the heterogeneous presentation of MDD, investigations

of more homogenous patient populations would both improve the interpretation of findings at the individual study level and promote more meaningful investigations at the meta-analytic level.

Convergence from Imaging Modalities

A central finding from the present study is the failure of convergence in the three single-modality meta-analyses and successful convergence – albeit relatively weak -- in pooled, multi-modality datasets. To our knowledge this is the largest meta-analysis in MDD to date and the first to pool results across imaging modalities (VBM and VBP) in this manner. Identification of convergent brain abnormalities across structural and functional dataset supports our hypothesis for the co-localization of disease effects in MDD. This co-localization, and/or longitudinal progression, of MDD-specific disease effects are not well established in current literature and call for further investigation. As the meta-analyses here were greatly facilitated by access to the collated and coded VBM literature shared in the BrainMap database, we anticipate that expanding BrainMap to include a sector sharing the resting-state VBP literature will be an important tool for future meta-analyses.

Overall, convergent findings in the present meta-analysis are sparse compared to the volume of input data (see SF1). Though the identification of regions of significant convergence in the present study is a distinct advance from previous meta-analyses – which failed to yield any convergent findings – the sparseness of present results is nonetheless notable. In our largest analysis (All Effects) only 8 of 102 total experiments contributed to the identified region of convergence. In our smallest analysis, (VBP-neg, Drug Naive) only 3 of 20 total experiments contributed to the identified region (see ST12 for details). The small number of contributing experiments is a stark contrast to the large

volume of findings reported at the individual study level. An interpretation of sparse findings from the present study may be attributed to our analysis of exclusively voxel-wise whole-brain studies. Conclusions from Sprooten et al's 2017 meta-analysis (which included analysis of both whole-brain and ROI-based studies) found that though ROI-based approaches seem more adept at yielding significant findings, this may be due to confirmation bias stemming from *a priori* selection of brain regions for selective analysis. As a result of this bias, Sprooten reported that there may be an artificial exaggeration of particular brain regions' role in psychiatric diseases and findings from ROI-based studies should be interpreted with caution. Considering the broad neuroimaging literature, the scarcity of results at the meta-analytic level underscores the impact of ROI-driven findings (16), clinical heterogeneity (6), and overall replication problem in current literature.

Finally, our findings suggest the possibility of improved convergence in task-independent data over task-activation data. The lack of convergence in the 2017 task-activation meta-analysis conducted by our group, Müller et al, was speculated to be due, in part, to confounds introduced through inconsistency of the processes investigated in various tasks (6). Further pitfalls of task-activation based studies include the dependence of task-based paradigms on patient cooperation (86) and lack of diagnostic specificity in findings from task-fMRI studies (16). The advantages of task-free paradigms, especially for use in meta-analysis, however, are not definitively addressed in current literature and warrants further investigation.

Limitations

A primary motivation for this meta-analysis was to compare task-independent findings to those from Müller et al (our group, 6), which utilized task-activation data. Thus following the All Effects analysis, we performed post-hoc analysis (without correction for multiple

comparison) utilizing the same parameters from Müller for significance threshold and design for sub-meta-analyses (cluster-level inference of $P<0.05$ with a cluster-forming threshold of $P<0.001$). We have also reported the clusters that prevail after Bonferroni correction (cluster-level inference restricted to $P<0.0027$) in Table 2. The survival of two regions identified from Drug Naïve and Only MDD subgroups further supports our conclusion that sample homogeneity in MDD plays a major role in convergence of neuroimaging findings.

Another limitation of the present work is the inability of current CBMA methods to integrate findings from functional and effective connectivity studies. Functional and effective connectivity studies represent a rich corpus of literature: 36 were identified in the literature search for the present work. However, the majority of these (26 of 36) were regionally restricted, using ROIs for analysis or seeding, making them ineligible for ALE CBMA. Only 2 resting-state functional connectivity studies were whole-brain and voxel-wise. These were not included in the present analysis.

Finally, as previously discussed, recruitment of clinically heterogeneous MDD populations at the individual study level substantially contributed to the limitations of the present work. Whole-group analyses were limited by varied medication status and other factors relating to clinical heterogeneity of MDD, which we tested for in subgroup meta-analyses to the best of our ability. Fifty-one percent (51%) of all publications included in the present analysis recruited mixed MDD patient populations and pooled all patients into heterogeneous groups regardless of age of onset, disease duration/severity, and number of previous episodes. As discussed in our Methods and Results, the testing of subgroups based on MDD severity (first episode, chronic/recurrent, treatment resistant) would have provided more clinically meaningful findings. As such, our findings from clinical subgroups limited to medication status may not definitively indicate neurobiologically homogenous patient characteristics and could, instead, be due to other

methodological factors that we were not able to reliably test for. However, delineating patient groups to the best of our ability did improve convergence of results and indicates that clinical heterogeneity of MDD warrants further investigation in future neuroimaging studies.

Conclusions

Our findings suggest that MDD exhibits a concordance of abnormality in both structure (VBM) and function (VBP) in select brain regions. Our findings suggest the presence of MDD-associated brain features, in contrast to lack of disease-specific findings from previous transdiagnostic and MDD-specific meta-analyses. Per our successful integration of VBP findings, we recommend the addition of a VBP sector to the BrainMap database to facilitate future meta-analyses in this area of study. Finally, our analysis of clinical heterogeneity within this meta-analysis suggests that diverse patient populations may pose significant confounds in the neuroimaging findings in MDD.

References

1. World Health Organization. Depression and other common mental disorders: global health estimates [Internet]. 2017 [cited November 12, 2018]. Available from: https://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/
2. Holtzheimer PE. Future prospects in depression research. Dialogues Clin Neurosci 2006; 8(2):175-189
3. Masdeu JC. Neuroimaging in Psychiatric Disorders. Neurotherapeutics 2011; 8(1):93-102
4. Kupfer DJ, Frank E, Phillips ML. Major depressive disorder: new clinical, neurobiological, and treatment perspectives. Lancet 2012; 379(9820):1045-1055
5. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson E, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013; 145(5):365-376
6. Müller VI, Cieslik EC, Serbanescu I, Laird AR, Fox PT, Eickhoff SB. Altered Brain Activity in Unipolar Depression Revisited: Meta-analyses of Neuroimaging Studies. JAMA Psychiatry 2017; 74(1):47–55
7. Ecklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci U S A 2016; 113(28):7900-7905
8. Fox PT, Lancaster JT, Laird AR, Eickhoff SB. Meta-Analysis in Human Neuroimaging: Computational Modeling of Large-Scale Databases. Annu Rev Neurosci 2014; 37:409-434

9. Samartsidis P, Montagna S, Nichols TE, Johnson TD. The coordinate-based meta-analysis of neuroimaging data. Stat Sci 2017; 32(4):580-599
10. Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. Activation likelihood estimation meta-analysis revisited. Neuroimage 2012; 59(3):2349-2361
11. Ashberner J, Friston KJ. Voxel-based morphometry-- the methods. Neuroimage 2000; 11(6 Pt 1):805-821
12. Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. Meta-analysis of the Functional Neuroanatomy of Single-Word Reading: Method and Validation. Neuroimage 2002; 16(3 Pt A):765-780
13. Laird AR, Fox PM, Price CJ, Glahn DC, Uecker AM, Lancaster JL, et al. ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. Hum Brain Mapp 2005; 25(1):155-164
14. Turkeltaub PE, Eickhoff SB, Laird AR, Fox PM, Wiener M, Fox PT. Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analysis. Hum Brain Mapp 2011; 33(1):1-13
15. Acar F, Seurinck R, Eickhoff SB, Moerkerke B. Assessing robustness against potential publication bias in Activation Likelihood Estimation (ALE) meta-analysis for fMRI. PLoS One 2018; 13(11):e0208177
16. Sprooten E, Rasgon A, Goodman M, Carlin A, Leibu E, Lee WH, et al. Addressing reverse inference in psychiatric neuroimaging: Meta-analyses of task-related brain activation in common mental disorders. Hum Brain Mapp 2017; 38(4):1846-1864
17. Goodkind M, Eickhoff SB, Oathes DJ, Jiang Y, Chang A, Jones-Hagata LB, et al. Identification of a common neurobiological substrate for mental illness. JAMA Psychiatry 2015; 72(4):305-315

18. Vanasse T, Fox PM, Barron D, Robertson M, Eickhoff SB, Lancaster JL, et al. BrainMap VBM: An Environment for Structural Meta-analysis. Hum Brain Mapp 2018; 39(8):3308-3325
19. Sha Z, Xia M, Lin Q, Cao M, Tang Y, Xu K, et al. Meta-Connectomic Analysis Reveals Commonly Disrupted Functional Architectures in Network Modules and Connectors across Brain Disorders. Cereb Cortex 2018; 28(12):4179-4194
20. Eickhoff SB, Nichols TE, Laird AR, Hoffstaedter F, Amunts K, Fox PT, et al. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 2016; 137:70-85
21. Müller VI, Cieslik EC, Laird AR, Fox PT, Radua J, Mataix-Cols D, et al. Ten simple rules for neuroimaging meta-analysis. Neurosci Biobehav Rev 2018; 84:151-161
22. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010;167(7):748–751
23. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Medicine 2013; 11(126)
24. González-Redondo R, García-García D, Clavero P, Gasca-Salas C, García-Eulate R, Zubieta JL, et al. Grey matter hypometabolism and atrophy in Parkinson's disease with cognitive impairment: a two-step process. Brain 2014; 137(Pt 8):2356-2367
25. Jack JR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer’s pathological cascade. Lancet Neurol 2010; 9(1):119-128

26. Botha H, Utianski RL, Whitwell JL, Duffy JR, Clark HM, Strand EA, et al. Disrupted functional connectivity in primary progressive apraxia of speech. Neuroimage Clin 2018; 18:617-629
27. Sanchis-Segura C, Cruz-Gómez AJ, Belenguer A, Fittipaldi Márguez MS, Ávila C, Forn C. Increased regional gray matter atrophy and enhanced functional connectivity in male multiple sclerosis patients. Neurosci Lett 2016; 630:154-157
28. Baltruschat SA, Ventura-Campos N, Cruz-Gómez ÁJ, Belenguer A, Forn C. Gray matter atrophy is associated with functional connectivity reorganization during the Paced Auditory Serial Addition Test (PASAT) execution in Multiple Sclerosis (MS). J Neuroradiol 2015; 42(3)
29. Lawrie SM, McIntosh AM, Hall J, Owens DG, Johnstone EC. Brain Structure and Function Changes During the Development of Schizophrenia: The Evidence From Studies of Subjects at Increased Genetic Risk. Schizophr Bull 2008; 34(2):330-340
30. Drevets WC, Price JL, Simpson JR, Todd RD, Reich T, Vannier M, et al. Subgenual prefrontal cortex abnormalities in mood disorders. Nature 1997; 386(6627);824-827
31. Drevets WC, Price JL, Furey ML. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. Brain Struct Funct 2008; 213(1-2):93-118
32. Wang W, Zhao Y, Hu X, Huang X, Kuang W, Lui S, et al. Conjoint and dissociated structural and functional abnormalities in first-episode drug-naive patients with major depressive disorder: a multimodal meta-analysis. Sci Rep 2017; 7:10401

33. Sacher J, Neumann J, Fünfstück T, Soliman A, Villringer A, Schroeter ML. Mapping the depressed brain: a meta-analysis of structural and functional alterations in major depressive disorder. J Affect Disord 2012; 140(2):142-148
34. Ying W. A new hypothesis of neurodegenerative diseases: The deleterious network hypothesis. Med Hypotheses 1996; 47(4);307-313
35. Zhou J, Gennatas ED, Kramer JH, Miller BL, Seeley WW. Predicting Regional Neurodegeneration from the Healthy Brain Functional Connectome. *Neuron* 2012; 73(6):1216–1227
36. Tahmasian M, Shao J, Meng C, Grimmer T, Diehl-Schmid J, Yousefi BH, et al. Network Degeneration Hypothesis: Separating Individual Patients with Different Neurodegenerative Syndromes in a Preliminary Hybrid PET/MR Study. J Nucl Med 2016; 57(3):410-415
37. Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci 2011; 15(10):483-506
38. Palazidou E. The neurobiology of depression. Br Med Bull 2012; 101:127-145
39. Ye M, Yang T, Qing P, Lei X, Qiu J, Liu G. Changes of Functional Brain Networks in Major Depressive Disorder: A Graph Theoretical Analysis of Resting-State fMRI. PLoS One 2015; 10(9):e0133775
40. Li BJ, Friston K, Mody M, Wang HN, Lu HB, Hu DW. A brain network model for depression: From symptom understanding to disease intervention. CNS Neurosci Ther 2018; 24(11):1004-1019
41. Crossley NA, Mechelli A, Scott J, Carletti F, Fox PT, McGuire P, Bullmore ET. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. Brain 2014; 137(Pt 8):2382-2395
42. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. Neuron 2009; 62(1):42-52

43. Fox PT, Lancaster JL. Mapping context and content: The BrainMap model. Nature Rev Neurosci 2002; 3:319-321
44. Fox PT, Laird AR, Fox SP, Fox PM, Uecker AM, Crank M, et al. BrainMap taxonomy of experimental design: Description and evaluation. Hum Brain Mapp 2005; 25:185-198
45. Laird AR, Lancaster JL, Fox PT. BrainMap: The social evolution of a functional neuroimaging database. Neuroinformatics 2005; 3:65-78
46. Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, et al. Automated Talairach Atlas labels for functional brain mapping. Hum Brain Mapp 2000; 10:120-131
47. Barron DS, Fox MP, Laird AR, Robinson JL, Fox PT. Thalamic medial dorsal nucleus atrophy in medial temporal lobe epilepsy: A VBM meta-analysis. Neuroimage Clin 2013; 2:25-32
48. GingerALE [Software edition 3.0]. 2019. Retrieved from http://brainmap.org/ale/
49. Glahn DC, Laird AR, Ellison-Wright E, Thelen SM, Robinson JL, Lancaster JL, et al. Meta-Analysis of Gray Matter Anomalies in Schizophrenia: Application of Anatomic Likelihood Estimation and Network Analysis. Biol Psychiatry 2008; 64(9):774-781
50. Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Human Brain Mapp 2009. 30(9): 2907-2926
51. Tahmasian M, Sepehry AA, Samea F, Khodadadifar T, Soltaninejad Z, Javaheripour N, et al. Practical recommendations to conduct a neuroimaging meta-analysis for neuropsychiatric disorders. Hum Brain Mapp 2019; doi 10.1002

52. Wager TD, Lindquist MA, Nichols TE, Kober H, Snellenberg JX. Evaluating the consistency and specificity of neuroimaging data using meta-analysis. Neuroimage 2009; 45(1):S210-S221
53. Samartsidis P, Montagna S, Laird AR, Fox PT, Johnson TD, Nichols TE. Estimating the number of missing experiments in a neuroimaging meta-analysis. bioRxiv 2017
54. Mayberg HS, Liotti M, Brannan SK, McGinnis S, Mahurin RK, Jerabek PA, et al. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am J Psychiatry 1999; 156(5):675-682
55. Seminowicz DA, Mayberg HS, McIntosh AR, Goldapple K, Kennedy S, Segal Z, et al. Limbic-frontal circuitry in major depression: a path modeling metanalysis. Neuroimage 2004; 22(1):409-418
56. Hamani C, Mayberg H, Stone S, Laxton A, Haber S, Lozano AM. The subcallosal cingulate gyrus in the context of major depression. Biol Psychiatry 2011; 69(4):301-108
57. Gotlib IH, Sivers H, Gabrieli JD, Whitfield-Gabrieli S, Goldin P, Minor KL, et al. Subgenual anterior cingulate activation to valenced emotional stimuli in major depression. Neuroreport 2005; 16(16):1731-1734
58. Huebl J, Brücke C, Merkl A, Bajbouj M, Schneider GH, Kühn AA. Processing of emotional stimuli is reflected by modulations of beta band activity in the subgenual anterior cingulate cortex in patients with treatment resistant depression. Soc Cogn Affect Neurosci 2016; 11(8):1290-1298
59. Ramirez-Mahaluf JP, Perramon J, Otal B, Villoslada P, Compte A. Subgenual anterior cingulate cortex controls sadness-induced modulations of cognitive and emotional network hubs. Sci Rep 2018; 8(1):8566

60. Mayberg H, Lozano AM, Voon V, McNeely HE, Seminowicz D, Hamani C, Schwalb JM, Kennedy SH. Deep brain stimulation for treatment-resistant depression. Neuron 2005; 45(5):651-660
61. Johansen-Berg H, Gutman DA, Behrens TE, Matthews PM, Rushworth MF, Katz E, et al. Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. Cereb Cortex 2008; 18(6):1375-1383
62. Lozano AM, Mayberg HS, Giacobbe P, Hamani C, Craddock RC, Kennedy SH. Subcallosal cingulate gyrus deep brain stimulation for treatment-resistant depression. Biol Psychiatry 2008; 64(6):461-467
63. Lozano AM, Giacobbe P, Hamani C, Rizvi SJ, Kennedy SH, Kolivakis TT, et al. A multicenter pilot study of subcallosal cingulate area deep brain stimulation for treatment-resistant depression. J Neurosurg 2012; 116(2):315-322
64. Naurshima K, McCormick LM, Yamada T, Thatcher RW, Robinson RG. Subgenual cingulate theta activity predicts treatment response of repetitive transcranial magnetic stimulation in participants with vascular depression. J Neuropsychiatry Clin Neurosci 2010; 22(1):75-84
65. Fox MD, Buckner RL, White MP, Greicius MD, Pascual-Leone A. Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. Biol Psychiatry 2012; 72(7):595-603
66. Hadas I, Sun Y, Lioumis P, Zomorrodi R, Jones B, Voineskos D, et al. Association of Repetitive Transcranial Magnetic Stimulation Treatment With Subgenual Cingulate Hyperactivity in Patients With Major Depressive Disorder: A Secondary Analysis of a Randomized Clinical Trial. JAMA Netw Open 2019; 2(6):e195578

67. Singh A, Erwin-Grabner T, Sutcliffe G, Antal A, Paulus W, Goya-Maldonado R. Personalized repetitive transcranial magnetic stimulation temporarily alters default mode network in healthy subjects. Sci Rep 2019; 9(1):5631
68. Sheline YI, Wang PW, Gado MH, Csernansky JG, Vannier MW. Hippocampal atrophy in recurrent major depression. Proc Natl Acad Sci USA 1996; 93(9):3908-3913
69. Bremner JD, Narayan M, Anderson ER, Staib LH,Miller HL, Charney DS. hippocampal volume reduction in major depression. Am J Psychiatry 2000; 157(1):115-118
70. Campbell S, MacQueen G. The role of the hippocampus in the pathophysiology of major depression. J Psychiatry Neurosci 2004; 29(6):417-426
71. Sheline YI. Depression and the Hippocampus: Cause or Effect? Biol Psychiatry 2011; 70(4):308-309
72. MacQueen G, Frodl T. The hippocampus in major depression: evidence for the convergence of the bench and bedside in psychiatric research? Mol Psychiatry 2011; 16(3):252-264
73. Sampath D, Sathyanesan M, Newton SS. Cognitive dysfunction in major depression and Alzheimer's disease is associated with hippocampal–prefrontal cortex dysconnectivity. Neuropsychiatr Dis Treat 2017; 13:1506-1519
74. MacQueen GM, Galway TM, Hay J, Young LT, Joffe RT. Recollection memory deficits in patients with major depressive disorder predicted by past depressions but not current mood state or treatment status. Psychol Med 2002; 32(2):251-258
75. Dillon DG. The neuroscience of positive memory deficits in depression. Front Psychol 2015; 6:1295

76. 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(Pt 5):1498-1507
77. Rodríguez-Cano E, Sarró S, Monté GC, Maristany T, Salvador R, McKenna PJ, et al. Evidence for structural and functional abnormality in the subgenual anterior cingulate cortex in major depressive disorder. Psychol Med 2014; 44(15):3263-327
78. Victor TA, Furey ML, Fromm SJ, Ohman A, Drevets WC. Relationship Between Amygdala Responses to Masked Faces and Mood State and Treatment in Major Depressive Disorder. Arch Gen Psychiatry 2010; 67(11):1128-1138
79. Hamilton JP, Siemer M, Gotlib IH. Amygdala volume in Major Depressive Disorder: A meta-analysis of magnetic resonance imaging studies. Mol Psychiatry 2008; 13(11):993-1000
80. Lu Y, Liang H, Han D, Mo Y, Li Z, Cheng Y, et al. The volumetric and shape changes of the putamen and thalamus in first episode, untreated major depressive disorder. Neuroimage Clin 2016; 11:658-666
81. Keedwell PA, Andrew C, Williams SC, Brammer MJ, Phillips ML. The Neural Correlates of Anhedonia in Major Depressive Disorder. Bil Psychiatry 2005; 58(11):843-853
82. Cheng W, Rolls ET, Qiu J, Xie X, Wei D, Huang CC, et al. Increased functional connectivity of the posterior cingulate cortex with the lateral orbitofrontal cortex in depression. Transl Psychiatry 2018; 8(1):90
83. Bonte FJ, Trivedi MH, Devous MD Sr, Harris TS, Payne JK, Weinberg WA, et al. Occipital Brain Perfusion Deficits in Children with Major Depressive Disorder. J Nucl Med 2011; 42(7):1059-1061

84. Yue Y, Yuan Y, Hou Z, Jiang W, Bai F, Zhang Z. Abnormal Functional Connectivity of Amygdala in Late-Onset Depression Was Associated with Cognitive Deficits. PLoS One 2013; 8(9):e75058
85. Schmall L, Hibar DP, Sämann PG, Hall GB, Baune BT, Jahanshad N, et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. Mol Psychiatry 2017; 22(6):900-909
86. Shimony JS, Zhang D, Johnston JM, Fox MD, Roy A, Leuthardt EC. Resting State Spontaneous Fluctuations in Brain Activity: A New Paradigm for Presurgical Planning using fMRI. Acad Radiol 2009; 16(5):578-583

Figure Legends:

Figure 1- Study selection delineation diagram

A literature search identified an overall dataset of 92 independent publications reporting 152 experiments. *Note that 5 publications included investigations of both structural and functional changes, contributing a total of 97 “studies”. Experiment modalities included: regional cerebral blood flow (rCBF) studies utilizing single photon emission tomography (SPECT), arterial spin labeling (ASL), and oxygen-15 positron emission tomography (O-15 PET); regional homogeneity (ReHo); (frequency) amplitude of low frequency fluctuations (ALFF/fALFF); fluorodeoxyglucose metabolism (FDG-PET); and voxel-based morphometry investigations of gray matter volume.. Experiment classes included: decreased gray matter volume relative to controls (VBMneg), decreased and increased function relative to controls (VBPneg and VBPpos, respectively).

Figure 2- Meta-Analytic Groups Tested. First, a preliminary All Effects analysis of unified results across all imaging modalities was preformed. Results from all experiments were concatenated into unified disease-control contrast groups of VBMneg+VBPpos+VBPneg. Next,18 different meta-analyses were performed including the five major meta-analytic classes of decreased gray matter volume relative to controls (VBMneg), decreased (VBPneg) and increased (VBPos) function relative to controls; each rsFX group pooled with the GMV group (VBPneg + VBMneg and VBPpos + VBMneg), and subgroups of each meta-analytic class comprised of specific clinical populations where number of qualifying experiments exceeded 17.

Figure 3- Abnormal regions identified. Abnormal regions identified from All Effects analysis and meta-analytic groupings of: combined VBPpos and VBMneg, and VBPneg. Clinical subgroupings which also identified abnormal regions are shown.