Structural and functional neuroimaging of late-life depression: a coordinate-based meta-analysis

Amin Saberi[a], Esmaeil Mohammadi[b], Mojtaba Zarei[a], Simon B. Eickhoff[c,d] and Masoud Tahmasian [*a]

[a]*Institute of Medical Science and Technology, Shahid Beheshti University, Tehran, Iran*
[b]*Non-Communicable Diseases Research Center, Endocrinology and Metabolism Population Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran*
[c]*Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, Jülich, Germany*
[d]*Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany*

This is a post-peer-review, pre-copyedit version of an article published in Brain Imaging and Behavior. The final authenticated version is available online at: http://dx.doi.org/10.1007/s11682-021-00494-9

Abstract

Several neuroimaging studies have investigated localized aberrations in brain structure, function or connectivity in late-life depression, but the ensuing results are equivocal and often conflicting. Here, we provide a quantitative consolidation of neuroimaging in late-life depression using coordinate-based meta-analysis by searching multiple databases up to March 2020. Our search revealed 3252 unique records, among which we identified 32 eligible whole-brain neuroimaging publications comparing 674 patients with 568 controls. The peak coordinates of group comparisons between the patients and the controls were extracted and then analyzed using activation likelihood estimation method. Our sufficiently powered analysis on all the experiments, and more homogenous subsections of the data (patients > controls, controls > patients, and functional imaging experiments) revealed no significant convergent regional abnormality in late-life depression. This inconsistency might be due clinical and biological heterogeneity of LLD, as well as experimental (e.g., choice of tasks, image modalities) and analytic flexibility (e.g., preprocessing and analytic parameters), and distributed patterns of neural abnormalities. Our findings highlight the importance of clinical/biological heterogeneity of late-life depression, in addition to the need for more reproducible research by using pre-registered and standardized protocols on more homogenous populations to identify potential consistent brain abnormalities in late-life depression.

Keywords: Late-life depression; Activation likelihood estimation; Functional magnetic resonance imaging; Voxel-based morphometry; Positron emission tomography.

1 Introduction

Late-life depression (LLD) is defined as major depressive disorder (MDD) in patients over the age of 50 (Vaishnavi & Taylor, 2006). This definition includes both late-onset depression (LOD), in which depression has started in later life, and early-onset depression (EOD), in which depression is first diagnosed in early adulthood and continues into older ages. In the year 2017 the global number of individuals with MDD who were older than 50 was estimated at about 60 million. In the general population and among the age groups of 50-69 and over 70, the prevalence of MDD was estimated at 3.3% and 3.7%, respectively, causing 2.5% and 3.8% of the total years lost due to disability (YLD) in these age groups. With the aging of the populations, the global number of LLD patients has increased by 27.1% from 2007 to 2017 (Global Burden of Disease Collaborative Network, 2018). LLD has detrimental effects on the mental well-being of the patients, and can lead to emotional impairment, cognitive dysfunction, and medical problems. It is a heterogeneous neuropsychiatric syndrome with variable presentations, including depressed mood, anxiety, psychomotor retardation, fatigue, feelings of guilt, hopelessness, worthlessness, and restricted mental, physical or social functioning (Nelson, Clary, Leon, & Schneider, 2005; Rutherford, Taylor, Brown, Sneed, & Roose, 2017; Szanto et al., 2012). In addition, cognitive dysfunction occurs in 20-25% of the patients and may involve impaired executive functioning, information processing and concentration, as well as explicit learning and memory (Elderkin-Thompson, Moody, Knowlton, Hellemann, & Kumar, 2011; Koenig, Bhalla, & Butters, 2014; Lamar, Charlton, Zhang, & Kumar, 2012). The cognitive

dysfunction may further progress and in some cases precedes Alzheimer's disease, which has an increased risk of 65% in LLD patients (Butters et al., 2008; Diniz, Butters, Albert, Dew, & Reynolds, 2013). In addition, LLD may be associated with amyloid beta accumulation (Mahgoub & Alexopoulos, 2016; Wu et al., 2014), although some studies have found normal or decreased amyloid beta accumulations in these patients (De Winter et al., 2017; Mackin et al., 2021). Furthermore, LLD increases the risk of developing/exacerbation of chronic medical diseases such as diabetes mellitus, cardiovascular diseases, and arthritis (Karakus & Patton, 2011; Zivin, Wharton, & Rostant, 2013). Older adults with depression are also at a greater risk of mortality, both from suicide, and by an increased rate of cardiovascular diseases (Wei et al., 2019). Despite its detrimental effects, depression in older adults is underdiagnosed and poorly responds to treatment, highlighting the need for more clear understanding of its neurobiology (Manning, Wang, & Steffens, 2019).

Aiming to unravel the neurobiological mechanisms of LLD, several neuroimaging studies have investigated localized abnormalities of the brain structure, function and connectivity, using voxel-based morphometry (VBM), task-based and resting-state functional magnetic resonance imaging (t-fMRI and rs-fMRI), and positron emission tomography (PET). Although these studies have advanced our understanding on the neural correlates of LLD, they have often reported conflicting and heterogeneous results. For example, structural neuroimaging studies have variably reported cortical atrophy or hypertrophy of cortical regions such as the orbitofrontal cortex, precuneus, lateral temporal, cingulate or insula (Byun et al., 2016; Harada et al., 2016; Hwang et al., 2010; Oudega et al., 2014; Smith, Kramer, et al., 2009). In addition, functional neuroimaging studies have found abnormalities, including increased or decreased functional activation or connectivity, in various structures such as the superior and inferior frontal gyri, precuneus, precentral gyrus, cingulate gyrus, parahippocampal cortex, cerebellum, or putamen (Bricenõ et al., 2015; Dombrovski, Szanto, Clark, Reynolds, & Siegle, 2013; Liu et al., 2012; Smith, Kramer, et al., 2009; Yuan, Zhang, et al., 2008).

In the context of these inconsistent findings, meta-analytic approaches are valuable tools for quantitatively consolidating the effects observed across the published literature in order to identify convergent findings (Müller et al., 2018; Tahmasian et al., 2019). Activation likelihood estimation (ALE) is a state-of-art coordinate-based meta-analysis (CBMA) method that identifies spatial convergence (or lack thereof) across findings of individual studies, and distinguishes true convergence from random overlap (Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012).

Of note, recent CBMAs on MDD have failed to determine convergence in their primary analyses, but revealed effects for restricted sub-analyses of the data, which may hint at the data heterogeneity (Gray, Müller, Eickhoff, & Fox, 2020; Müller et al., 2017). And indeed, psychopathologically, LLD differs from depression in early adulthood e.g., by increased prevalence of cognitive dysfunction and somatic symptoms, but less dominant sadness feeling (Fiske, Wetherell, & Gatz, 2009; Hegeman, Kok, van der Mast, & Giltay, 2012), and a weaker response to the antidepressants (Tedeschini et al., 2011). In addition, imaging findings such as hippocampal atrophy, or microvascular lesions, are more common/severe in LLD patients (Aizenstein et al., 2016; McKinnon, Yucel, Nazarov, & MacQueen, 2009). Therefore, and based on these differences between mid- and late-life depression, a neuroimaging meta-analysis restricted to LLD is much needed to identify convergent neuroimaging findings specific to this age group. Here, we used the ALE method on the reported brain differences of LLD patients and healthy individuals derived from the whole-brain structural and functional neuroimaging studies, to provide a quantitative assessment of convergence across published literature.

2 Methods

This study was pre-registered on the International Prospective Register of Systematic Reviews (PROSPERO, code: CRD42019115872), and is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). Here, we followed the most recent best-practice guidelines for neuroimaging meta-analyses (Müller et al., 2018; Tahmasian et al., 2019), to perform an ALE meta-analysis on the neuroimaging studies comparing LLD patients with healthy control (HC) participants.

2.1 Search and study selection

We searched PubMed, Embase, Scopus, and Web of Science databases in March 2020 using the following search terms: (elderly OR geriatric OR "late life" OR "later life" OR "late onset" OR older OR "old age") AND (depress* OR MDD) AND ("voxel based morphometry" OR VBM OR "functional magnetic resonance imaging" OR fMRI OR "Positron-Emission Tomography" OR PET). The detailed search strategy for each database is reported in Table S1. In addition to the records obtained by the search, we traced the references of relevant neuroimaging reviews/meta-analyses and the included studies. Af-

ter removing records duplicated in multiple databases, a total of 3253 unique records were screened by two independent reviewers (A.S., E.M.). The screening was performed in two stages, first using titles and/or abstracts, and then using full texts of the potentially relevant records identified in the first stage.

We included original studies in which: (1) depression was diagnosed through an interview and using standardized diagnostic criteria (i.e. Diagnostic and Statistical Manual of Mental Disorders [DSM], or International Classification of Diseases [ICD]), (2) patients had no major psychological or neurological comorbidities, such as dementia, stroke, Parkinson's disease, or psychosis (although comorbid anxiety was allowed due to its high co-occurrence with depression [Beekman et al., 2000]), (3) patients were compared to elderly healthy control individuals, (4) all participants in both groups were > 50 years old, (5) whole-brain structural or functional grey matter differences were assessed using rs-fMRI, t-fMRI, VBM, or PET, (6) analysis was not limited to a region of interest (ROI), and small volume correction (SVC) was not performed, as these approaches are biased toward finding significance in the respective regions, hence violating the assumption of ALE method that all voxels of the brain have a unified chance of being reported (Müller et al., 2018; Tahmasian et al., 2019), (7) patients were not part of an interventional study, unless a baseline comparison with healthy controls was reported, (8) peak coordinates of significant findings were reported in Montreal Neurological Institute (MNI) or Talairach standard spaces, or were provided by authors upon our request. Of note, we excluded the studies with no significant findings (Colloby et al., 2011; Delaloye et al., 2010; Marano et al., 2015; Patel et al., 2012; Sexton et al., 2012; Sin et al., 2018; Smith, Reynolds, et al., 2009; Vanyukov et al., 2015; Weber et al., 2010, 2012), as ALE method is aimed at identifying spatial convergence of findings, for which these studies contributed no data.

2.2 Data extraction

The extracted data consisted of bibliographic information (first author, title, journal, country and institute), demographic and clinical data (number of participants, age, sex, age of onset, clinical status, medication status), methodological details (imaging modality, scanner field strength, task name and domain, software package, analysis approach, covariates, method of multiple comparison correction), and the peak coordinates of between-group experiments reported in each study. Notice that here, "study" refers to an individual publication, whereas "experiment" refers to individual contrasts reported within a "study" (e.g. LLD > HC, and HC > LLD), each yielding a distinct set of coordinates. Coordinates reported in Talairach space were transformed into MNI space (Lancaster et al., 2007), so that all the experiments are in the same reference space. If the applied reference space was not explicitly reported or provided by authors at our request, as suggested, we estimated it from the default settings of the software packages used for normalization - i.e., FSL, SPM and FreeSurfer use MNI, while BrainVoyager uses Talairach (Müller et al., 2018; Tahmasian et al., 2019).

In ALE meta-analyses, pooling the data from overlapping samples causes spurious findings by improperly increasing the influence of that sample (Turkeltaub et al., 2012). Therefore, we took great care to avoid convergence over analyses performed on (partially) overlapping samples, both within and across papers. We reviewed the included studies for signs of overlap with other studies, by examining their team members, location of study, recruitment interval, and sample demographics, and merged their data in such cases. For the same reason, in both primary and complementary ALE analyses (see below), all the coordinates from multiple experiments (i.e., LLD > HC or HC > LLD, different imaging modalities, or different tasks) pertaining to the same subjects were merged, to make sure that in all analyses each sample is only represented by one experiment.

2.3 Activation likelihood estimation

The revised version of the ALE method (Eickhoff et al., 2012) was used to test the spatial convergence of the reported differences, against the null hypothesis of randomly distributed findings across the brain. In this method, the peak coordinates, or foci, are convolved with 3D Gaussian probability distributions that have a full width at half maximum (FWHM) inversely proportional to the sample size. This allows experiments with larger samples to have a greater statistical certainty in the meta-analysis. Next, for each experiment, the convolved foci are combined to generate per-experiment "modeled activation" (MA) maps. Subsequently, the MA maps for all the experiments included in the meta-analysis are combined into an ALE score map, representing the convergence of results at each particular location of the brain. The ALE score map is then statistically tested against a null distribution reflecting randomly distributed findings, to distinguish true convergence from by-chance overlap (Eickhoff et al., 2012; Turkeltaub et al., 2012). Finally, to avoid spurious findings, the resulting p-values are corrected for multiple comparison using the stringent family-wise error correction at the cluster level (cFWE), thresholded at $p < 0.05$ (Eickhoff, Laird, Fox, Lancaster, & Fox, 2017). All these procedures were implemented using an in-house MATLAB script.

We performed separate ALE meta-analyses on differ-

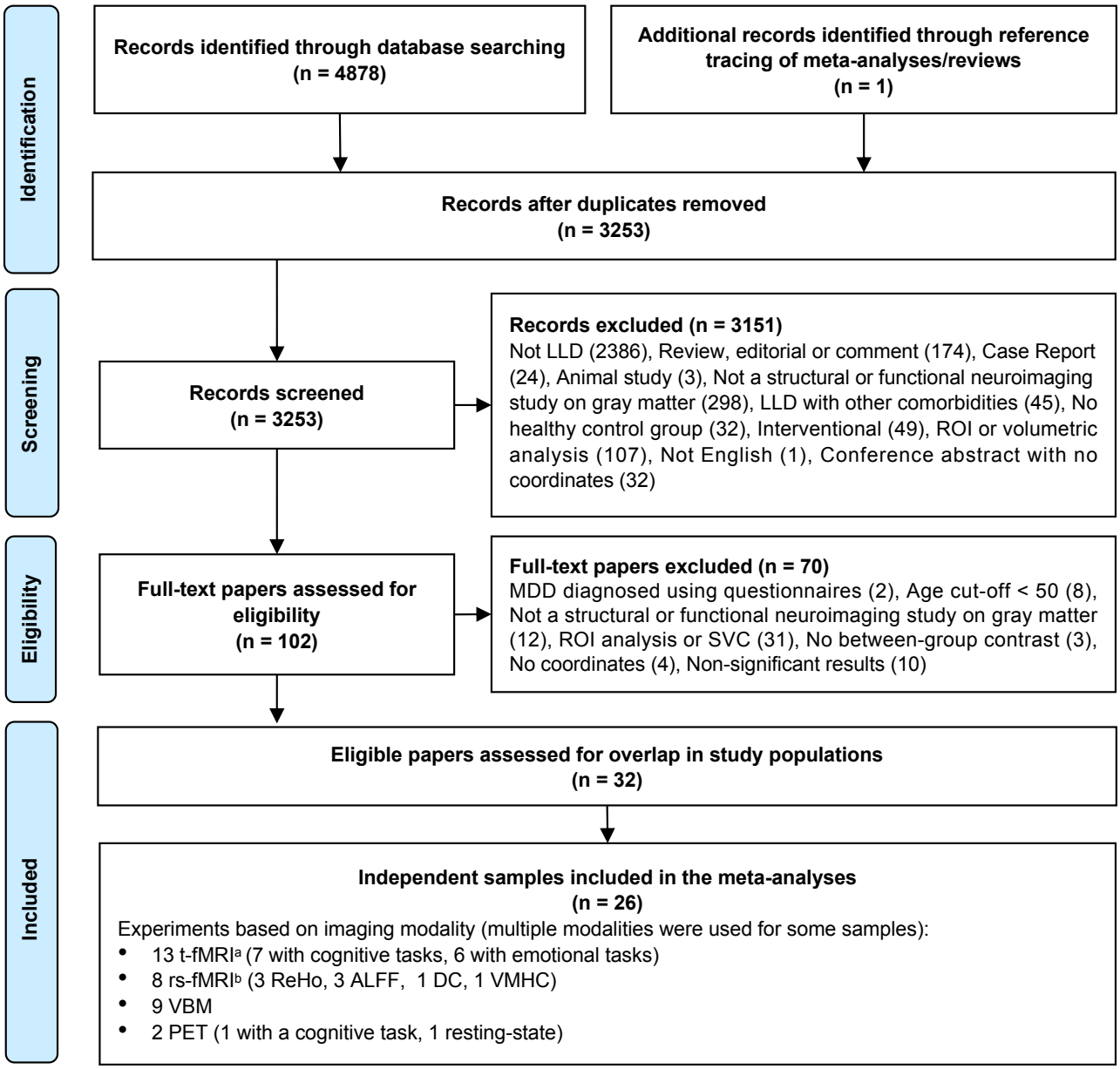

Figure 1: Study selection flowchart.

LLD: late-life depression; MDD: major depressive disorder; ROI: region of interest; SVC: small volume correction; rs-fMRI: resting-state functional magnetic resonance imaging; t-fMRI: task-based functional magnetic resonance imaging; ReHo: regional homogeneity; ALFF: amplitude of low frequency fluctuations; DC: voxel-wise degree centrality; VMHC: voxel-mirrored homotopic connectivity; VBM: voxel-based morphometry; PET: positron emission tomography. [a] Two different types of cognitive tasks and one emotional task were used for the same sample. [b] Both ReHo and ALFF were used in the same sample.

ent sets of experiments. In the primary analysis, to investigate all the brain abnormalities, including increased or decreased gray matter density, functional activation, or functional connectivity, we assessed the convergence across all the experiments. Next, we split the experiments by the direction of effect (HC > LLD or LLD > HC), imaging modalities (VBM, rs-fMRI, t-fMRI, or PET), task domains (cognitive vs. emotional), age of onset (LOD vs. EOD), or disease status (acutely depressed vs. patients in remission), and performed complementary analyses on more homogeneous subsections of the data that fulfilled the requirement of representing at least 15 experiments for sufficient power (Eickhoff et al., 2016).

2.4 ROI-based analysis

In addition to the whole-brain analysis, we investigated the convergence of the reported coordinates within the 7 cortical resting-state networks defined by (Yeo et al., 2011) which includes the visual, somatomotor, dorsal attention, salience / ventral attention, limbic, control and default networks. For each network, we used a permutation test to compare the sum of all ALE values in the network to their distribution under null, which tests for the maximum convergence for the average across the entire network. The p-values in this analysis were corrected for multiple comparisons using Bonferroni's correction (= 0.05 / 7 networks = 0.007).

3 Results

We identified 32 eligible whole-brain neuroimaging papers comparing LLD patients with HC individuals (Albert, Gau, Taylor, & Newhouse, 2017; Bobb et al., 2012; Bricenõ et al., 2015; Byun et al., 2016; Chen et al., 2012; De Asis et al., 2001; Dombrovski et al., 2013; Dumas & Newhouse, 2015; Fang et al., 2015; Guo et al., 2013; Harada et al., 2018, 2016; Hou, Sui, Song, & Yuan, 2016; C. M. Huang et al., 2019; Hwang et al., 2010; Lee, Liu, Wai, Ko, & Lee, 2013; Li et al., 2020; Liu et al., 2012; Mah, Williams, Leung, Freel, & Pollock, 2011; Oudega et al., 2014; Rao et al., 2015; Respino et al., 2019; Ribeiz et al., 2013; Smith, Kramer, et al., 2009; Takami, Okamoto, Yamashita, Okada, & Yamawaki, 2007; L. Wang et al., 2008; Weisenbach et al., 2014; Wong et al., 2016; Xie et al., 2012; Yuan, Zhang, et al., 2008; Yuan, Zhu, et al., 2008; Yue, Jia, Hou, Zang, & Yuan, 2015). Among these, four groups of papers had (partially) overlapping samples, including (Bricenõ et al., 2015; Rao et al., 2015; Weisenbach et al., 2014), (Chen et al., 2012; Guo et al., 2013; Liu et al., 2012), (Harada et al., 2018, 2016), and (Yuan, Zhang, et al., 2008; Yuan, Zhu, et al., 2008). The data from these papers was merged, yielding 26 independent study populations with 674 LLD and 568 HC participants (Fig. 1).

VBM was used in nine experiments, t-fMRI in 11 experiments, rs-fMRI in seven experiments, and PET in two experiments. Task-based fMRI experiments were performed with various cognitive (N = 7) and emotional (N = 6) tasks. Of note, one t-fMRI study employed two different cognitive tasks and one emotional task, which were reported in separate papers (Bricenõ et al., 2015; Rao et al., 2015; Weisenbach et al., 2014). The demographic, clinical, and technical characteristics of the included studies are summarized in Tables 1 and S2.

Fig. 2 illustrates the spatial distribution for peak coordinates of all the experiments based on the direction of the effects. Neither the primary nor the supplementary analyses revealed any significant regional convergence of neuroimaging findings for LLD (Table 2). In the primary analysis, 26 experiments showing increased or decreased grey matter density, functional activation, or functional connectivity of the brain in LLD patients yielded pcFWE = 0.828. Pooling over 17 experiments reflecting increases (i.e. LLD > HC), and 19 experiments representing decreases (i.e. HC > LLD) separately, resulted in pcFWE = 0.181 and pcFWE = 0.903, respectively. Restricting the analysis to experiments using functional neuroimaging (N = 20), or fMRI (N = 18), likewise, resulted in no significant convergence, with pcFWE = 0.544 and pcFWE = 0.409, respectively. Repeating all analyses with threshold-free cluster enhancement (TFCE), which is a potentially more lenient method of multiple comparison correction, again resulted in no significant convergence. Furthermore, ROI-based analysis revealed no convergence of the reported coordinates within the visual (punc = 0.650), somatomotor (punc = 0.459), dorsal attention (punc = 0.962), salience / ventral attention (punc = 0.049), limbic (punc = 0.675), control (punc = 0.635), and default (punc = 0.216) networks after Bonferroni's correction.

4 Discussions

Following the best-practice guidelines for conducting meta-analyses, we observed no significant convergence of regional brain abnormalities in LLD in both primary and complementary analyses. The observed lack of convergence indicates that the current literature on LLD does not support consistent, localized pathophysiology. Aspects that may have contributed to this null effect will be discussed as follows.

4.1 Clinical and biological heterogeneity of LLD

Depression is a heterogeneous diagnostic category with regard to the individual differences in presentations, treatment outcomes, comorbidities, genetic etiologies and biological mechanisms (Goldberg, 2011; Lynch, Gunning, & Liston, 2020). Considering the symptoms profile alone, more than 200 unique combinations of symptoms can theoretically fulfil the DSM criteria for MDD (Lynch et al., 2020). In addition, evidence from biochemical, genetics, and neuroimaging studies have suggested that MDD is biologically heterogeneous, meaning that different mechanistic pathways converge to a common phenotype, i.e., depression (for review see Beijers, Wardenaar, van Loo, & Schoevers, 2019). For instance, using resting-state functional connectivity and symptom profiles, four subtypes of depression have been identified

	First Author (Year)a	Number (% female)		Age mean ± SD / median (IQR)		Acute Depresssion	Anti-depressants	Age of Onset	Neuroimaging Technique	Task
		LLD	HC	LLD	HC					
1	Albert, K (2017)	12 (100%)	21 (100%)	62.4 ± 5.7	60.6 ± 6.8	No	Variable	n.r.	t-fMRI	emotion dot probe
2	Bobb, DS (2012)	15 (80%)	13 (69%)	60.7 ± 4.7	62.0 ± 5.3	Yes	Off	n.r.	t-fMRI	stop signal task
	Bricenō, EM (2015)	26 (46%)	25 (48%)	65.4 ± 8.1	68.1 ± 8.2				t-fMRI	facial emotion perception
3	Rao, JA (2015)	24 (n.r.)	23 (n.r.)	66.8 ± 8.2	67.9 ± 8.1	Yes	Variable	EOD	t-fMRI	go/no-go
	Weisenbach, SL (2014)	24 (58%)	23 (43%)	66.8 ± n.r.	67.9 ± n.r.				t-fMRI	semantic list learning
4	Byun, MS (2016)	29 (72%)	27 (48%)	71.6 ± 5.0	68.7 ± 6.0	Variable	On	LOD	VBM	
	Chen, JD (2012)	16 (60%)	15 (60%)	67.5 ± 6.1	64.9 ± 3.7				rs-fMRI (ReHo)	
5	Guo, W (2013)	17 (60%)	16 (60%)	67.5 ± 6.1	64.9 ± 3.7	Yes	Drug-naive	LOD	rs-fMRI (ALFF)	
	Liu, F (2012)	15 (60%)	15 (60%)	67.5 ± 6.1	64.9 ± 3.7				rs-fMRI (ReHo)	
6	De Asis, J (2001)	6 (0%)	5 (0%)	70.7 ± n.r.	67.6 ± n.r.	Yes	Variable	n.r.	PET	
7	Dombrovski, AY (2013)	31 (58%)	20 (60%)	66.3 ± 5.8	70.7 ± 8.7	Variable	Variable	n.r.	t-fMRI	probabilistic reversal learning
8	Dumas, JA (2015)	11 (n.r.)	12 (n.r.)	73.3 ± 3.3	72.6 ± 5.7	n.r.	Variable	Variable	t-fMRI	n-back
9	Fang, J (2015)	20 (00%)	18 (0%)	59.2 + 3.7	59.1 ± 7.5	Variable	n.r.	n.r.	VBM, rs-fMRI (ALFF)	
10	Harada, K (2016)	45 (67%)	61 (57%)	60.2 ± 8.2	62.9 ± 7.6	Yes	Variable	n.r.	VBM	
	Harada, K (2018)	16 (63%)	30 (62%)	56 (53.5-65.5)	58 (54-67)				VBM	
11	Hou, Z (2016)	31 (67%)	37 (51%)	68.0 ± 6.0	65.2 ± 7.5	Yes	Off	LOD	rs-fMRI (VMHC)	
12	Huang, CM (2019)	55 (69%)	40 (62%)	66.3 ± 5.4	68.1 ± 5.3	Yes	Variable	LOD	t-fMRI	color-word emotional Stroop
13	Hwang, JP (2010)	70 (00%)	26 (0%)	79.4 ± 5.3	79.5 ± 4.3	Yes	n.r.	LOD	VBM	
14	Lee, TW (2013)	14 (21%)	14 (35%)	65.1 ± 4.9	64.8 ± 4.2	Yes	On	LOD	t-fMRI	n-back
15	Li, J (2020)	50 (62%)	33 (48%)	66.6 ± 0.7	67.2 ± 0.8	Yes	Variable	Variable	rs-fMRI (DC)	
16	Mah, L (2011)	5 (0%)	8 (0%)	66 ± 6	69 ± 5	n.r.	Drug-naive	n.r.	t-fMRI	emotional judgment of faces
17	Oudega, ML (2014)	55 (65%)	23 (52%)	72.3 ± 7.8	70.3 ± 6.3	Yes	Off	Variable	VBM	
18	Respino, M (2019)	33 (63%)	43 (58%)	72.2 ± 6.6	73.4 ± 6.5	Yes	Off	n.r.	rs-fMRI	
19	Ribeiz, SRI (2013)	30 (76%)	22 (77%)	70.7 ± 6.5	70.4 ± 7.5	n.r.	n.r.	Variable	VBM	
20	Smith, GS (2009)	16 (62%)	13 (61%)	65.3 ± 9.1	67.4 ± 7.4	n.r.	Off	n.r.	VBM, PET	
21	Takami, H (2007)	10 (70%)	10 (60%)	62.5 ± 9.1	67.6 ± 9.7	No	Variable	LOD	t-fMRI	word generation
22	Wang, L (2008)	27 (44%)	20 (60%)	70.0 ± 5.7	73.1 ± 5.1	Variable	Variable	LOD	t-fMRI	emotional oddball task
23	Wong, NM (2016)	31 (54%)	23 (60%)	67.4 ± 5.4	67.1 ± 4.7	n.r.	On	LOD	t-fMRI	emotion processing
24	Xie, C (2012)	18 (77%)	25 (48%)	68.6 ± 6.8	74.2 ± 8.2	Yes	Variable	n.r.	VBM	
25	Yuan, Y (2008)	18 (55%)	14 (50%)	67.2 ± 7.3	67.1 ± 4.8	No	Off	LOD	rs-fMRI (ReHo)	
	Yuan, Y (2008)	19 (52%)	16 (50%)	67.2 ± 7.3	67.1 ± 4.8				VBM	
26	Yue, Y (2015)	16 (50%)	16 (50%)	68.1 ± 5.2	68.2 ± 4.6	Yes	Drug-naive	LOD	rs-fMRI (ALFF)	

Table 1: Characteristics of studies included in the meta-analysis.

LLD = late-life derpession; HC = healthy control; SD = standard deviaiton; n.r. = not reported; EOD = early-onset depression; LOD = late-onset depression; t-fMRI = task-based functional magnetic resonance imaging; rs-fMRI = resting-state functional magnetic resonance imaging; VBM = voxel-based morphometry; PET = positron-emission tomography; ReHo = regional homogeneity; ALFF = amplitude of low frequency fluctuations; DC = voxel-wise degree centrality; VMHC = voxel-mirrored homotopic connectivity. [a] Publications with overlapping samples are grouped together

Imaging modality	Contrast	Number of experiments	p-value	
			TFCE	cFWE
Structural or functional (VBM, PET, fMRI)	-	26	0.568	0.828
	HC >LLD	19	0.755	0.903
	LLD >HC	17	0.136	0.181
Functional (PET or fMRI)	-	20	0.216	0.544
fMRI (task-based or resting-state)	-	18	0.183	0.409

Table 2: Results of ALE meta-analyses on patients with late-life depression compared to healthy subjects by modality and direction of effect.

VBM = voxel-based morphometry; PET = positron-emission tomography; fMRI = functional magnetic resonance imaging; LLD = late-life depression; HC = healthy control; TFCE = threshold-free cluster enhancement; cFWE = family-wise error at cluster level.

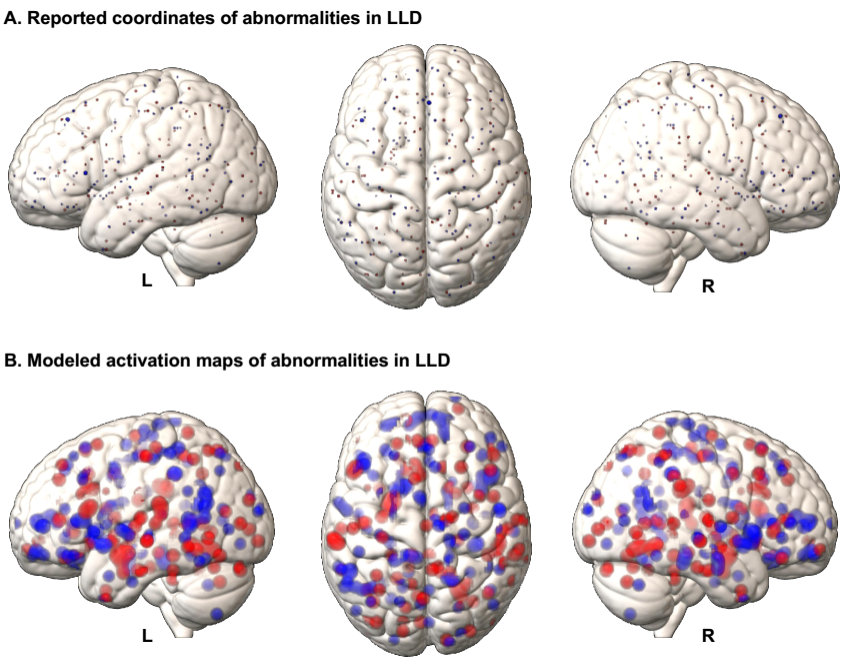

Figure 2: The activation likelihood estimation meta-analysis steps.

Panel A shows the distribution of reported coordinates reflecting structural/functional brain alterations in patients with late-life depression (LLD) compared to healthy control (HC) subjects. Panel B displays the modeled activation maps, obtained by combining the reported coordinates modeled as 3D Gaussian distributions (red = LLD > HC, blue = HC > LLD). Statistical testing against null-distribution, corrected for multiple comparisons using family-wise error correction at the cluster level, showed no convergence of the reported abnormalities.

with distinct patterns of dysfunctional connectivity in limbic and frontostriatal networks, distinct clinical profiles, and varying treatment responsiveness to repetitive transcranial magnetic stimulation (Drysdale et al., 2017).

In the older adults, due to the aging and aging-related comorbidities, the clinical and biological heterogeneity of depression is more prominent (Alexopoulos, 2019; Rutherford et al., 2017). Using clinical variables, a data-driven analysis identified five distinct subtypes of LLD, including 'mild pure depression', 'severe pure depression', 'amnestic depression', 'frail-depressed, physically dominated', and 'frail-depressed, cognitively dominated' with distinct remission and mortality rates (Lugtenburg et al., 2020). Biologically, variable phenotypes of LLD may result from multiple interacting or independent pathways, including cerebrovascular aging, inflammation, and oxidative stress (reviewed by Alexopoulos, 2019; Rutherford et al., 2017). Cerebrovascular aging and the ensuing small infarcts, which appear as white matter hyperintensities (WMHs), may damage deep white matter tracts (Li Wang, Leonards, Sterzer, & Ebinger, 2014). These infarcts can disrupt the connectivity of prefrontal cortical regions to the limbic and striatal areas (Wen, Steffens, Chen, & Zainal, 2014), leading to a subtype of LLD characterized by executive dysfunction and anhedonia (Alexopoulos, 2019). Inflammation may be the main culprit in another subtype of patients, in which dopaminergic functioning is disturbed due to inflammation (Bäckman, Nyberg, Lindenberger, Li, & Farde, 2006) and leads to cognitive and motor retardation and impaired level of activity (Rutherford et al., 2017). Oxidative stress can also lead to depression through an alternative mechanism, in which damages to mitochondrial DNA impairs its function, and in turn decreases the amount of energy available, causing fatigue, reduced physical and mental activity, and frailty (Rutherford et al., 2017). Cognitive dysfunction is another important source of heterogeneity in LLD. This symptom occurs only in a subset of LLD patients (20-25%) (Koenig et al., 2014), is presumed to be caused by variable mechanisms, including glucocorticoid-induced hippocampal atrophy, cerebrovascular aging, or Alzheimer's disease-related pathologies, and has variable clinical courses, with some remaining stable, but others progressing to vascular dementia or Alzheimer's disease (Butters et al., 2008; Koenig et al., 2014). Differences in the medication status, clinical status (actue or chronic), severity of disease, comorbidities, sex and age of onset additionally

contribute to the heterogeneity of LLD. For example, LOD and EOD are different in many aspects, such as clinical symptoms, genetic susceptibility, microvascular abnormalities, and neuroimaging findings (Tittmann et al., 2014). Taken together, LLD is a clinically and biologically heterogeneous disease, likely caused by variable pathophysiological processes, which are further confounded by the presence or absence of other aging-related comorbidities and neurodegenerative processes. Therefore, we can argue that our lack of convergent findings could in part be attributed to the varying combinations of subtypes included in each study and the ensuing dilution of subtype-specific effects at the level of individual studies and the meta-analysis.

4.2 Beyond the localized abnormalities of grey matter

LLD may be associated with distributed brain network abnormalities, rather than localized ones. In this case, neuroimaging studies aimed at localizing the effects would each pick up on different (due to noise and flexibility, cf. section 4.1) distinct parts of the disturbed, distributed network, resulting in poor convergence of local findings (Kharabian Masouleh, Eickhoff, Hoffstaedter, Genon, & Alzheimer's Disease Neuroimaging Initiative, 2019). Of note, in ALE meta-analyses, all seed-based connectivity studies are excluded in order to ensure that all the voxels have the same a priori chance of being reported (Müller et al., 2018; Tahmasian et al., 2019). In addition, ALE meta-analysis is a method of finding localized convergent abnormalities in the brain structure, function or connectivity, which does not test the network alterations associated with LLD. Nevertheless, we used a ROI-based ALE approach to show that the reported foci are not preferentially convergent in any of the specific resting-state brain networks. The functional connectivity studies on LLD have been reviewed elsewhere (Manning et al., 2019; Tadayonnejad & Ajilore, 2014), indicating increased functional connectivity in the default mode network and salience network, but decreased functional connectivity in executive control network, in addition to decreased network efficiency, reduced network strength, and increased long-range connections in LLD patients, which are similar to the functional connectivity changes in mid-life depression (Ebneabbasi et al., 2021; Mulders, van Eijndhoven, Schene, Beckmann, & Tendolkar, 2015).

In addition, beyond the grey matter, the pathophysiology of LLD is also associated with micro-/macrostructural white-matter abnormalities. As described above, cerebrovascular aging and WMHs are associated with an increased risk of LLD (van Agtmaal, Houben, Pouwer, Stehouwer, & Schram, 2017), particularly when located in the cingulum, uncinate fasciculus, superior longitudinal fasciculus, frontal lobe, and temporal lobe (Alexopoulos, 2019). In addition, a meta-analysis on diffusion tensor imaging (DTI) studies of LLD reported lower fractional anisotropy of the dorsolateral prefrontal cortex and uncinate fasciculus, indicating a possible role for the disruption of frontal and frontal-to-limbic white matter tracts in the pathogenesis of LLD (Wen et al., 2014).

4.3 Heterogeneity of the neuroimaging methodology

The studies included in this meta-analysis used a wide range of imaging acquisition techniques, preprocessing, and analytic methods to explore various neurobiological features of LLD. One notable source of methodological heterogeneity across the included studies was the different applied structural and functional imaging modalities including VBM, rs-fMRI, t-fMRI, or PET. In addition, further within the functional imaging studies (i.e., rs-fMRI, t-fMRI, or PET), the subjects were engaged in different mental states, either in the resting state (30%), or while doing a variety of cognitive (30%) or emotional (23%) task paradigms (e.g., n-back, stop-signal, Go/No-Go, word generation, emotional judgement of faces) with different visual or auditory stimuli which naturally involve distinct neural processes. We can appreciate how this can influence the findings, by looking at the results of three included studies that using different tasks (i.e., facial emotion perception, Go/No-Go, and semantic list learning) on the same sample found different, and in some regions conflicting, results (Bricenõ et al., 2015; Rao et al., 2015; Weisenbach et al., 2014). In rs-fMRI experiments, an additional source of heterogeneity may be the various applied analytical approaches, each designed to capture a conceptually different aspect of the whole-brain functional connectivity. For instance, while Amplitude of Low Frequency Fluctuations (ALFF; 11%) assesses the regional intensity of oscillatory fluctuations in a voxel's time series (Zou et al., 2008), Regional Homogeneity (ReHo; 11%) calculates the correlation of a voxel's time series with that of its nearest voxels (Zang, Jiang, Lu, He, & Tian, 2004), and voxel-wise Degree Centrality (DC; 3%) is a graph theory measure that represents the total weights of connections for a given voxel (Rubinov & Sporns, 2010).

In both structural and functional neuroimaging studies, preprocessing steps and analytical flexibility regarding design matrices, multiple comparison correction methods, software packages, and even operating systems introduces another level of heterogeneity across studies (Glatard et al., 2015; Poldrack et al., 2017). Of note, the methodological flexibility of neuroimaging studies is so diverse, that in a recent study, among 70

independent teams analyzing the same fMRI dataset, not even two teams chose identical analytic workflows (Botvinik-Nezer et al., 2020).The experimental and analytical flexibility of neuroimaging methods is a double-edge sword, as it allows investigating many diverse aspects of neurobiology, but can also lead to spurious/false positive findings. This is particularly true for low-powered studies and in the presence of liberal thresholding and selective reporting (Button et al., 2013). Low-powered studies are inherently prone to an increased proportion of false positive findings, i.e., small positive predictive value (Ioannidis, 2005). In addition, they are more likely to fail to identify the effects, particularly when they are small, and in a research community that values statistical significance, are incentivized to adopt questionable research practices, or "p-hacking", to "fish" for significant findings (Button et al., 2013). For instance, they might choose to report the results that are rather liberal and uncorrected for multiple comparisons, as in 31% of our included studies. In this case, and particularly given that the sample sizes in our included studies were in the low to moderate range (median: 45, range: 11-96), some of the reported foci included in our meta-analysis may be non-replicable.

4.4 Overview of previous neuroimaging meta-analyses on LLD

Previous CBMAs on LLD have solely focused on the structural brain abnormalities, and reported conflicting results (Boccia, Acierno, & Piccardi, 2015; Du et al., 2014). They have identified convergence in a variety of regions including the putamen, hippocampus, amygdala, parahippocampal gyrus, medial frontal gyrus, subcallosal gyrus, and lingual gyrus (Du et al., 2014), as well as the amygdala, thalamus, cingulate gyrus, precuneus, superior frontal gyrus, and ventromedial frontal cortex (Boccia et al., 2015). According to the recent best-practice guidelines for neuroimaging meta-analyses, there are some methodological issues in these studies. Firstly, they lacked sufficient statistical power, as only nine studies (including four studies with null results) were included in (Du et al., 2014), and only six studies (nine experiments) were included in (Boccia et al., 2015), both below the number required for a robust meta-analysis (15), which makes their results unstable and susceptible to the influence of a single experiment (Eickhoff et al., 2016). Of note, the number of structural neuroimaging experiments included in our meta-analysis (N = 9) was not sufficient, and we forewent a separate analysis on these experiments. Secondly, in (Du et al., 2014) effect size signed differential mapping (ES-SDM) was utilized, which is an alternative CBMA method, and considered statistically more lenient compared to ALE, and is also different by taking studies with null findings into account. Thirdly, both meta-analyses used liberal methods of multiple comparison correction, and were susceptible to false-positive findings. Specifically, (Du et al., 2014) reported the results with uncorrected threshold of $p = 0.005$ and $Z > 1$, and (Boccia et al., 2015) used false discovery rate (FDR), which is more liberal compared to cFWE, and is no longer recommended in ALE meta-analyses (Eickhoff et al., 2016; Müller et al., 2018; Tahmasian et al., 2019).

4.5 Limitations, recommendations and future directions

ALE analyses with few number of experiments are unstable and largely influenced by a single experiment (Eickhoff et al., 2016), and therefore, except for the complementary analyses limited to the functional experiments, our other pre-planned subgroup analyses limited to e.g. structural experiments, rs-fMRI experiments, t-fMRI experiments, specific task domains/paradigms, specific methods of rs-fMRI analysis, or LOD/EOD were not viable. It is important to note that because of the differences between the imaging modalities, multimodal CBMAs are subject to debate. However, a trade-off is inherent to all meta-analytic approaches, including CBMAs, between the power or generalizability of the findings versus the homogeneity of the included experiments, with their goal often being to maximize power/generalizability while minimizing heterogeneity to a reasonable extent, which is defined by the research question of interest and the amount of available data (Tahmasian, Zarei, et al., 2018). Multimodal CBMAs, e.g. (Gray et al., 2020; Noordermeer, Luman, & Oosterlaan, 2016; Radua et al., 2012; Raschle, Menks, Fehlbaum, Tshomba, & Stadler, 2015; Samea et al., 2019; Tahmasian, Noori, et al., 2018), average over perhaps interesting modality-specific findings, but have the advantage of identifying robust findings which are consistent across a larger number and more diverse range of studies. Of note, with more neuroimaging studies published on LLD, future updates to this meta-analysis would be able to perform additional subgroup analyses, and identify convergence across more homogenous data. In addition, CBMAs are based on limited spatial data, i.e., the peak coordinates of significant regions between groups, and are hence less powerful compared to the alternative image-based meta-analyses, which consolidate the findings of studies using their unthresholded statistical parametric maps (Salimi-Khorshidi, Smith, Keltner, Wager, & Nichols, 2009). Here, due to the sparsity of the available statistical parametric maps, we were unable to do an image-based meta-analysis on LLD. However, we encourage authors to submit their statistical para-

metric maps to open data sharing platforms such as NeuroVault repository (https://neurovault.org/; Gorgolewski et al., 2015), so that this more powerful method of meta-analysis will be possible in the future. Finally, an additional limitation of our study was that due to a lack of consensus on the LLD definition, our included studies had arbitrarily used variable age cut-offs to define LLD (ranging from 50 to 65, but more commonly 60).

This study, and previous meta-analyses on MDD (Gray et al., 2020; Müller et al., 2017), have shown inconsistency of findings across neuroimaging studies on (late-life) depression. Of note, this is not limited to depression, and has also been observed in e.g. insomnia disorder, narcolepsy, Parkinson's disease, migraine, attention deficit hyperactivity disorder, self-injurious thoughts, or the impact of COMT Val158Met allele on brain activation related to working memory (Giehl, Tahmasian, Eickhoff, & van Eimeren, 2019; X. Huang, Rootes-Murdy, Bastidas, Nee, & Franklin, 2020; Nickl-Jockschat, Janouschek, Eickhoff, & Eickhoff, 2015; Samea et al., 2019; Sheng et al., 2020; Tahmasian, Noori, et al., 2018; Tench, Tanasescu, Cottam, Constantinescu, & Auer, 2019). In order to differentiate a true lack of localizable effects from between-study variability, we recommend future studies on (late-life) depression to (1) use larger sample sizes ideally through collaborations such as The Enhancing NeuroImaging Genetics through Meta-Analysis consortium (Schmaal et al., 2020; Thompson et al., 2020) to ameliorate site-idiosyncrasies, (2) openly share their data to allow replication and future integration, (3) standardize experimental, preprocessing, and analytical methods or at least establish replicability through providing sufficient detail or code (Button et al., 2013), (4) focus more on multivariate analytical approaches, which are better equipped to detect abnormal brain networks (Kharabian Masouleh et al., 2019), (5) consider data-driven approaches to accommodate potential biological subtypes, and (6) pre-register protocols for individual studies, as well for meta-analyses to reduce publication bias.

5 Conclusions

In summary, following the current best-practice protocols, in this pre-registered and sufficiently powered analyses we did not observe any consistent local abnormality in LLD. The combination of clinical/biological subtypes and distributed patterns, in addition to the heterogeneity of protocols and selective reporting may explain the lack of any localizable convergence for LLD. Our finding emphasizes the importance of identifying LLD subtypes, and taking them into account in the future studies and highlights the need for using standard methodology of future neuroimaging studies on (late-life) depression, toward more reproducible, pre-registered and clearly reported studies, which recruit more homogenous populations, and are aimed at identifying both localized and distributed abnormalities of the brain.

Acknowledgements

None.

Author Contributions

MT and AS formed the idea and designed the study. AS and EM contributed to the literature search and data extraction, based on suggestions from MT and SBE. AS and MT performed the analyses. SBE contributed algorithms and additional supervision. AS drafted the manuscript. MT, SBE, MZ, and EM revised and approved the paper.

Funding Sources

This research did not receive any funding.

Compliance with Ethical Standards

This study was a systematic review on the previously published studies and did not use original human or animal data.

Conflict of Interest

None of the authors have a conflict of interest to declare.

Data Availability

Data is available upon reasonable request.

References

Aizenstein, H. J., Baskys, A., Boldrini, M., Butters, M. A., Diniz, B. S., Jaiswal, M. K., … Tene, O. (2016). Vascular depression consensus report—A critical update. Bmc Medicine, 14(1), 161. doi: 10.1186/s12916-016-0720-5

Albert, K., Gau, V., Taylor, W. D., & Newhouse, P. A. (2017). Attention bias in older women with remitted depression is associated with enhanced amygdala activity and functional connectivity. Journal of Affective Disorders, 210, 49–56. doi: 10.1016/j.jad.2016.12.010

Alexopoulos, G. S. (2019). Mechanisms and treatment of late-life depression. Translational Psychiatry, 9(1), 188. doi: 10.1038/s41398-019-0514-6

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. Neuroscience and Biobehavioral Reviews, 30(6), 791–807. doi: 10.1016/j.neubiorev.2006.06.005

Beekman, A. T. F., de Beurs, E., van Balkom, A. J. L. M., Deeg, D. J. H., van Dyck, R., & van Tilburg, W. (2000). Anxiety and Depression in Later Life: Co-Occurrence and Communality of Risk Factors. American Journal of Psychiatry, 157(1), 89–95. doi: 10.1176/ajp.157.1.89

Beijers, L., Wardenaar, K. J., van Loo, H. M., & Schoevers, R. A. (2019). Data-driven biological subtypes of depression: Systematic review of biological approaches to depression subtyping. Mol Psychiatry, 24(6), 888–900. doi: 10.1038/s41380-019-0385-5

Bobb, D. S., Adinoff, B., Laken, S. J., McClintock, S. M., Rubia, K., Huang, H. W., … Andrew Kozel, F. (2012). Neural correlates of successful response inhibition in unmedicated patients with late-life depression. American Journal of Geriatric Psychiatry, 20(12), 1057–1069. doi: 10.1097/JGP.0b013e318235b728

Boccia, M., Acierno, M., & Piccardi, L. (2015). Neuroanatomy of Alzheimer's Disease and Late-Life Depression: A Coordinate-Based Meta-Analysis of MRI Studies. Journal of Alzheimer's Disease: JAD, 46(4), 963–970. doi: 10.3233/jad-142955

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. Nature, 582(7810), 84–88. doi: 10.1038/s41586-020-2314-9

Bricenõ, E. M., Rapport, L. J., Kassel, M. T., Bieliauskas, L. A., Zubieta, J. K., Weisenbach, S. L., & Langenecker, S. A. (2015). Age and gender modulate the neural circuitry supporting facial emotion processing in adults with major depressive disorder. American Journal of Geriatric Psychiatry, 23(3), 304–313. doi: 10.1016/j.jagp.2014.05.007

Butters, M. A., Young, J. B., Lopez, O., Aizenstein, H. J., Mulsant, B. H., Reynolds, C. F., … Becker, J. T. (2008). Pathways linking late-life depression to persistent cognitive impairment and dementia. Dialogues in Clinical Neuroscience, 10(3), 345–357. doi: 10.31887/DCNS.2008.10.3/mabutters

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. Nature Reviews. Neuroscience, 14(5), 365–376. doi: 10.1038/nrn3475

Byun, M. S., Choe, Y. M., Sohn, B. K., Yi, D., Han, J. Y., Park, J., … Lee, D. Y. (2016). Association of cerebral amyloidosis, blood pressure, and neuronal injury with late-life onset depression. Frontiers in Aging Neuroscience, 8(OCT). doi: 10.3389/fnagi.2016.00236

Chen, J. D., Liu, F., Xun, G. L., Chen, H. F., Hu, M. R., Guo, X. F., … Zhao, J. P. (2012). Early and late onset, first-episode, treatment-naive depression: Same clinical symptoms, different regional neural activities. Journal of Affective Disorders, 143(1–3), 56–63. doi: 10.1016/j.jad.2012.05.025

Colloby, S. J., Firbank, M. J., Vasudev, A., Parry, S. W., Thomas, A. J., & O'Brien, J. T. (2011). Cortical thickness and VBM-DARTEL in late-life depression. Journal of Affective Disorders, 133(1–2), 158–164. doi: 10.1016/j.jad.2011.04.010

De Asis, J. M., Stern, E., Alexopoulos, G. S., Pan, H., Van Gorp, W., Blumberg, H., … Silbersweig, D. A. (2001). Hippocampal and anterior cingulate activation deficits in patients with geriatric depression. American Journal of Psychiatry, 158(8), 1321–1323. doi: 10.1176/appi.ajp.158.8.1321

De Winter, F.-L., Emsell, L., Bouckaert, F., Claes, L., Jain, S., Farrar, G., … Vandenbulcke, M. (2017). No Association of Lower Hippocampal Volume With Alzheimer's Disease Pathology in Late-Life Depression. The American Journal of Psychiatry, 174(3), 237–245. doi: 10.1176/appi.ajp.2016.16030319

Delaloye, C., Moy, G., de Bilbao, F., Baudois, S., Weber, K., Hofer, F., … Gold, G. (2010). Neuroanatomical and neuropsychological features of elderly euthymic depressed patients with early- and late-onset. Journal of the Neurological Sciences, 299(1–2), 19–23. doi: 10.1016/j.jns.2010.08.046

Diniz, B. S., Butters, M. A., Albert, S. M., Dew, M. A., & Reynolds, C. F. (2013). Late-life depression and risk of vascular dementia and Alzheimer's disease: Systematic review and meta-analysis of community-based cohort studies. The British Journal of Psychiatry: The Journal of Mental Science, 202(5), 329–335. doi: 10.1192/bjp.bp.112.118307

Dombrovski, A. Y., Szanto, K., Clark, L., Reynolds, I. C. F., & Siegle, G. J. (2013). Reward signals, attempted suicide, and impulsivity in late-life depression. Jama Psychiatry, 70(10), 1020–1030. doi: 10.1001/jamapsychiatry.2013.75

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med, 23(1), 28–38. PubMed (27918562). doi: 10.1038/nm.4246

Du, M., Liu, J., Chen, Z., Huang, X., Li, J., Kuang, W., … Gong, Q. (2014). Brain grey matter volume alterations in late-life depression. Journal of Psychiatry and Neuroscience, 39(6), 397–406. doi: 10.1503/jpn.130275

Dumas, J. A., & Newhouse, P. A. (2015). Impaired working memory in geriatric depression: An fMRI study. American Journal of Geriatric Psychiatry, 23(4), 433–436. doi: 10.1016/j.jagp.2014.09.011

Ebneabbasi, A., Mahdipour, M., Nejati, V., Li, M., Liebe, T., Colic, L., … Tahmasian, M. (2021). Emotion processing and regulation in major depressive disorder: A 7T resting-state fMRI study. Human Brain Mapping, 42(3), 797–810. doi: 10.1002/hbm.25263

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. NeuroImage, 59(3), 2349–2361. doi: 10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2017). Implementation errors in the GingerALE Software: Description and recommendations. Human Brain Mapping, 38(1), 7–11. doi: 10.1002/hbm.23342

Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., … Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. NeuroImage, 137, 70–85. doi: 10.1016/j.neuroimage.2016.04.072

Elderkin-Thompson, V., Moody, T., Knowlton, B., Hellemann, G., & Kumar, A. (2011). Explicit and implicit memory in late-life depression. The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 19(4), 249–255. doi: 10.1097/JGP.0b013e3181e89a5b

Fang, J., Mao, N., Jiang, X., Li, X., Wang, B., & Wang, Q. (2015). Functional and Anatomical Brain Abnormalities and Effects of Antidepressant in Major Depressive Disorder: Combined Application of Voxel-Based Morphometry and Amplitude of Frequency Fluctuation in Resting State. Journal of Computer Assisted Tomography, 39(5), 766–773. doi: 10.1097/rct.0000000000000264

Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in older adults. Annu Rev Clin Psychol, 5, 363–389. doi: 10.1146/annurev.clinpsy.032408.153621

Giehl, K., Tahmasian, M., Eickhoff, S. B., & van Eimeren, T. (2019). Imaging executive functions in Parkinson's disease: An activation likelihood estimation meta-analysis. Parkinsonism & Related Disorders, 63, 137–142. doi: 10.1016/j.parkreldis.2019.02.015

Glatard, T., Lewis, L. B., Ferreira da Silva, R., Adalat, R., Beck, N., Lepage, C., … Evans, A. C. (2015). Reproducibility of neuroimaging analyses across operating systems. Frontiers in Neuroinformatics, 9. doi: 10.3389/fninf.2015.00012

Global Burden of Disease Collaborative Network. (2018). Global Burden of Disease Study 2017 (GBD 2017) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME). Available from http://ghdx.healthdata.org/gbd-results-tool.

Goldberg, D. (2011). The heterogeneity of "major depression." World Psychiatry, 10(3), 226–228. doi: 10.1002/j.2051-5545.2011.tb00061.x

Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., … Margulies, D. S. (2015). NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Frontiers in Neuroinformatics, 9. doi: 10.3389/fninf.2015.00008

Gray, J. P., Müller, V. I., Eickhoff, S. B., & Fox, P. T. (2020). Multimodal Abnormalities of Brain Structure and Function in Major Depressive Disorder: A Meta-Analysis of Neuroimaging Studies. The American Journal of Psychiatry, 177(5), 422–434. doi: 10.1176/appi.ajp.2019.19050560

Guo, W. B., Liu, F., Xun, G. L., Hu, M. R., Guo, X. F., Xiao, C. Q., … Zhao, J. P. (2013). Reversal alterations of amplitude of low-frequency fluctuations in early and late onset, first-episode, drug-naive depression. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 40(1), 153–159. doi: 10.1016/j.pnpbp.2012.08.014

Harada, K., Ikuta, T., Nakashima, M., Watanuki, T., Hirotsu, M., Matsubara, T., … Matsuo, K. (2018). Altered connectivity of the anterior cingulate and the posterior superior temporal gyrus in a longitudinal study of later-life depression. Frontiers in Aging Neuroscience, 10(FEB). doi: 10.3389/fnagi.2018.00031

Harada, K., Matsuo, K., Nakashima, M., Hobara, T., Higuchi, N., Higuchi, F., … Watanabe, Y. (2016). Disrupted orbitomedial prefrontal limbic network in individuals with later-life depression. Journal of Affective Disorders, 204, 112–119. doi: 10.1016/j.jad.2016.06.031

Hegeman, J. M., Kok, R. M., van der Mast, R. C., & Giltay, E. J. (2012). Phenomenology of depression in older compared with younger adults: Meta-analysis. Br J Psychiatry, 200(4), 275–281. doi: 10.1192/bjp.bp.111.095950

Hou, Z., Sui, Y., Song, X., & Yuan, Y. (2016). Disrupted interhemispheric synchrony in default mode network underlying the impairment of cognitive flexibility in late-onset depression. Frontiers in Aging Neuroscience, 8(SEP). doi: 10.3389/fnagi.2016.00230

Huang, C. M., Fan, Y. T., Lee, S. H., Liu, H. L., Chen, Y. L., Lin, C., & Lee, T. M. C. (2019). Cognitive reserve-mediated neural modulation of emotional control and regulation in people with late-life depression. Social Cognitive and Affective Neuroscience, 14(8), 849–860. doi: 10.1093/scan/nsz054

Huang, X., Rootes-Murdy, K., Bastidas, D. M., Nee, D. E., & Franklin, J. C. (2020). Brain Differences Associated with Self-Injurious Thoughts and Behaviors: A Meta-Analysis of Neuroimaging Studies. Scientific Reports, 10(1), 2404. doi: 10.1038/s41598-020-59490-6

Hwang, J. P., Lee, T. W., Tsai, S. J., Chen, T. J., Yang, C. H., Lirng, J. F., & Tsai, C. F. (2010). Cortical and subcortical abnormalities in late-onset depression with history of suicide attempts investigated with MRI and voxel-based morphometry. Journal of Geriatric Psychiatry and Neurology, 23(3), 171–184. doi: 10.1177/0891988710363713

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. PLOS Medicine, 2(8), e124. doi: 10.1371/journal.pmed.0020124

Karakus, M. C., & Patton, L. C. (2011). Depression and the onset of chronic illness in older adults: A 12-year prospective study. The Journal of Behavioral Health Services & Research, 38(3), 373–382. doi: 10.1007/s11414-011-9234-2

Kharabian Masouleh, S., Eickhoff, S. B., Hoffstaedter, F., Genon, S., & Alzheimer's Disease Neuroimaging Initiative. (2019). Empirical examination of the replicability of associations between brain structure and psychological variables. ELife, 8, e43464. doi: 10.7554/eLife.43464

Koenig, A. M., Bhalla, R. K., & Butters, M. A. (2014). Cognitive functioning and late-life depression. Journal of the International Neuropsychological Society, 20(5), 461–467. doi: 10.1017/S1355617714000198

Lamar, M., Charlton, R., Zhang, A., & Kumar, A. (2012). Differential associations between types of verbal memory and prefrontal brain structure in healthy aging and late life depression. Neuropsychologia, 50(8), 1823–1829. doi: 10.1016/j.neuropsychologia.2012.04.007

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., … Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Human Brain Mapping, 28(11), 1194–1205. doi: 10.1002/hbm.20345

Lee, T. W., Liu, H. L., Wai, Y. Y., Ko, H. J., & Lee, S. H. (2013). Abnormal neural activity in partially remitted late-onset depression: An fMRI study of one-back working memory task. Psychiatry Research - Neuroimaging, 213(2), 133–141. doi: 10.1016/j.pscychresns.2012.04.010

Li, J., Gong, H., Xu, H., Ding, Q., He, N., Huang, Y., … Zhan, S. (2020). Abnormal Voxel-Wise Degree Centrality in Patients With Late-Life Depression: A Resting-State Functional Magnetic Resonance Imaging Study. Frontiers in Psychiatry, 10. doi: 10.3389/fpsyt.2019.01024

Liu, F., Hu, M., Wang, S., Guo, W., Zhao, J., Li, J., … Chen, H. (2012). Abnormal regional spontaneous neural activity in first-episode, treatment-naive patients with late-life depression: A resting-state fMRI study. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 39(2), 326–331. doi: 10.1016/j.pnpbp.2012.07.004

Lugtenburg, A., Zuidersma, M., Wardenaar, K. J., Aprahamian, I., Rhebergen, D., Schoevers, R. A., & Oude Voshaar, R. C. (2020). Late-life related subtypes of depression—A data-driven approach on cognitive domains and physical frailty. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences. doi: 10.1093/gerona/glaa110

Lynch, C. J., Gunning, F. M., & Liston, C. (2020). Causes and Consequences of Diagnostic Heterogeneity in Depression: Paths to Discovering Novel Biological Depression Subtypes. Biol Psychiatry. (32171465). doi: 10.1016/j.biopsych.2020.01.012

Mackin, R. S., Insel, P. S., Landau, S., Bickford, D., Morin, R., Rhodes, E., … Alzheimer's Disease Neuroimaging Initiative and the ADNI Depression Project. (2021). Late-Life Depression Is Associated With Reduced Cortical Amyloid Burden: Findings From the Alzheimer's Disease Neuroimaging Initiative Depression Project. Biological Psychiatry, 89(8), 757–765. doi: 10.1016/j.biopsych.2020.06.017

Mah, L., Williams, L. J., Leung, R., Freel, S., & Pollock, B. G. (2011). Neural substrates of emotional processing in late-life depression. Biological Psychiatry, 69(9), 139S. doi: 10.1016/j.biopsych.2011.03.031

Mahgoub, N., & Alexopoulos, G. S. (2016). The Amyloid Hypothesis: Is There a Role for Anti-amyloid Treatment in Late-Life Depression? The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 24(3), 239–247. doi: 10.1016/j.jagp.2015.12.003

Manning, K., Wang, L., & Steffens, D. (2019). Recent advances in the use of imaging in psychiatry: Functional magnetic resonance imaging of large-scale brain networks in late-life depression. F1000Res, 8. doi: 10.12688/f1000research.17399.1

Marano, C. M., Workman, C. I., Lyman, C. H., Munro, C. A., Kraut, M. A., & Smith, G. S. (2015). Structural imaging in late-life depression: Association with mood and cognitive responses to antidepressant treatment. American Journal of Geriatric Psychiatry, 23(1), 4–12. doi: 10.1016/j.jagp.2013.10.001

McKinnon, M. C., Yucel, K., Nazarov, A., & MacQueen, G. M. (2009). A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. Journal of Psychiatry & Neuroscience: JPN, 34(1), 41–54.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med, 6(7), e1000097. doi: 10.1371/journal.pmed.1000097

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

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., … Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. Neuroscience & Biobehavioral Reviews, 84, 151–161. doi: 10.1016/j.neubiorev.2017.11.012

Müller, V. I., Cieslik, E. C., Serbanescu, I., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2017). Altered Brain Activity in Unipolar Depression Revisited: Meta-analyses of Neuroimaging Studies. JAMA Psychiatry, 74(1), 47–55. doi: 10.1001/jamapsychiatry.2016.2783

Nelson, J. C., Clary, C. M., Leon, A. C., & Schneider, L. S. (2005). Symptoms of late-life depression: Frequency and change during treatment. The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 13(6), 520–526. doi: 10.1176/appi.ajgp.13.6.520

Nickl-Jockschat, T., Janouschek, H., Eickhoff, S. B., & Eickhoff, C. R. (2015). Lack of Meta-Analytic Evidence for an Impact of COMT Val158Met Genotype on Brain Activation During Working Memory Tasks. Biological Psychiatry, 78(11), e43–e46. doi: 10.1016/j.biopsych.2015.02.030

Noordermeer, S. D. S., Luman, M., & Oosterlaan, J. (2016). A Systematic Review and Meta-analysis of Neuroimaging in Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD) Taking Attention-Deficit Hyperactivity Disorder (ADHD) Into Account. Neuropsychology Review, 26(1), 44–72. doi: 10.1007/s11065-015-9315-8

Oudega, M. L., van Exel, E., Stek, M. L., Wattjes, M. P., van der Flier, W. M., Comijs, H. C., … van den Heuvel, O. A. (2014). The structure of the geriatric depressed brain and response to electroconvulsive therapy. Psychiatry Research - Neuroimaging, 222(1–2), 1–9. doi: 10.1016/j.pscychresns.2014.03.002

Patel, M. J., Boada, F. E., Price, J. C., Sheu, L. K., Tudorascu, D. L., Reynolds, I. C. F., & Aizenstein, H. J. (2012). Association of small vessel ischemic white matter changes with BOLD fMRI imaging in the elderly. Psychiatry Research - Neuroimaging, 204(2–3), 117–122. doi: 10.1016/j.pscychresns.2012.09.006

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., … Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. Nature Reviews. Neuroscience, 18(2), 115–126. doi: 10.1038/nrn.2016.167

Radua, J., Borgwardt, S., Crescini, A., Mataix-Cols, D., Meyer-Lindenberg, A., McGuire, P. K., & Fusar-Poli, P. (2012). Multimodal meta-analysis of structural and functional brain changes in first episode psychosis and the effects of antipsychotic medication. Neuroscience and Biobehavioral Reviews, 36(10), 2325–2333. doi: 10.1016/j.neubiorev.2012.07.012

Rao, J. A., Kassel, M. T., Weldon, A. L., Avery, E. T., Briceno, E. M., Mann, M., … Weisenbach, S. L. (2015). The double burden of age and major depressive disorder on the cognitive control network. Psychology and Aging, 30(2), 475–485. doi: 10.1037/pag0000027

Raschle, N. M., Menks, W. M., Fehlbaum, L. V., Tshomba, E., & Stadler, C. (2015). Structural and Functional Alterations in Right Dorsomedial Prefrontal and Left Insular Cortex Co-Localize in Adolescents with Aggressive Behaviour: An ALE Meta-Analysis. PloS One, 10(9), e0136553. doi: 10.1371/journal.pone.0136553

Respino, M., Hoptman, M. J., Victoria, L. W., Alexopoulos, G. S., Solomonov, N., Stein, A. T., … Gunning, F. M. (2019). Cognitive Control Network Homogeneity and Executive Functions in Late-Life Depression. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. doi: 10.1016/j.bpsc.2019.10.013

Ribeiz, S. R. I., Duran, F., Oliveira, M. C., Bezerra, D., Castro, C. C., Steffens, D. C., … Bottino, C. M. C. (2013). Structural Brain Changes as Biomarkers and Outcome Predictors in Patients with Late-Life Depression: A CrossSectional and Prospective Study. Plos One, 8(11). doi: 10.1371/journal.pone.0080049

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52(3), 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Rutherford, B. R., Taylor, W. D., Brown, P. J., Sneed, J. R., & Roose, S. P. (2017). Biological Aging and the Future of Geriatric Psychiatry. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 72(3), 343–352. doi: 10.1093/gerona/glw241

Salimi-Khorshidi, G., Smith, S. M., Keltner, J. R., Wager, T. D., & Nichols, T. E. (2009). Meta-analysis of neuroimaging data: A comparison of image-based and coordinate-based pooling of studies. NeuroImage, 45(3), 810–823. doi: 10.1016/j.neuroimage.2008.12.039

Samea, F., Soluki, S., Nejati, V., Zarei, M., Cortese, S., Eickhoff, S. B., … Eickhoff, C. R. (2019). Brain alterations in children/adolescents with ADHD revisited: A neuroimaging meta-analysis of 96 structural and functional studies. Neuroscience and Biobehavioral Reviews, 100, 1–8. doi: 10.1016/j.neubiorev.2019.02.011

Schmaal, L., Pozzi, E., C Ho, T., van Velzen, L. S., Veer, I. M., Opel, N., … Veltman, D. J. (2020). ENIGMA MDD: Seven years of global neuroimaging studies of major depression through worldwide data sharing. Translational Psychiatry, 10(1), 172. doi: 10.1038/s41398-020-0842-6

Sexton, C. E., Allan, C. L., Le Masurier, M., McDermott, L. M., Kalu, U. G., Herrmann, L. L., … Ebmeier, K. P. (2012). Magnetic resonance imaging in late-life depression: Multimodal examination of network disruption. Archives of General Psychiatry, 69(7), 680–689. doi: 10.1001/archgenpsychiatry.2011.1862

Sheng, L., Zhao, P., Ma, H., Yuan, C., Zhong, J., Dai, Z., & Pan, P. (2020). A lack of consistent brain grey matter alterations in migraine. Brain, 143(6), e45–e45. doi: 10.1093/brain/awaa123

Sin, E. L. L., Liu, H. L., Lee, S. H., Huang, C. M., Wai, Y. Y., Chen, Y. L., … Lee, T. M. C. (2018). The relationships between brain structural changes and perceived loneliness in older adults suffering from late-life depression. International Journal of Geriatric Psychiatry, 33(4), 606–612. doi: 10.1002/gps.4831

Smith, G. S., Kramer, E., Ma, Y., Kingsley, P., Dhawan, V., Chaly, T., & Eidelberg, D. (2009). The functional neuroanatomy of geriatric depression. International Journal of Geriatric Psychiatry, 24(8), 798–808. doi: 10.1002/gps.2185

Smith, G. S., Reynolds, I. C. F., Houck, P. R., Dew, M. A., Ginsberg, J., Ma, Y., … Pollock, B. G. (2009). Cerebral glucose metabolic response to combined total sleep deprivation and antidepressant treatment in geriatric depression: A randomized, placebo-controlled study. Psychiatry Research - Neuroimaging, 171(1), 1–9. doi: 10.1016/j.pscychresns.2008.05.001

Szanto, K., Dombrovski, A. Y., Sahakian, B. J., Mulsant, B. H., Houck, P. R., Reynolds, C. F., & Clark, L. (2012). Social emotion recognition, social functioning, and attempted suicide in late-life depression. The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 20(3), 257–265. doi: 10.1097/JGP.0b013e31820eea0c

Tadayonnejad, R., & Ajilore, O. (2014). Brain network dysfunction in late-life depression: A literature review. Journal of Geriatric Psychiatry and Neurology, 27(1), 5–12. doi: 10.1177/0891988713516539

Tahmasian, M., Noori, K., Samea, F., Zarei, M., Spiegelhalder, K., Eickhoff, S. B., … Eickhoff, C. R. (2018). A lack of consistent brain alterations in insomnia disorder: An activation likelihood estimation meta-analysis. Sleep Medicine Reviews, 42, 111–118. doi: 10.1016/j.smrv.2018.07.004

Tahmasian, M., Sepehry, A. A., Samea, F., Khodadadifar, T., Soltaninejad, Z., Javaheripour, N., … Eickhoff, C. R. (2019). Practical recommendations to conduct a neuroimaging meta-analysis for neuropsychiatric disorders. Human Brain Mapping, 40(17), 5142–5154. doi: 10.1002/hbm.24746

Tahmasian, M., Zarei, M., Noori, K., Khazaie, H., Samea, F., Spiegelhalder, K., … Eickhoff, C. R. (2018). Reply to Hua Liu, HaiCun Shi and PingLei Pan: Coordinate based meta-analyses in a medium sized literature: Considerations, limitations and road ahead. Sleep Medicine Reviews, 42, 236–238. doi: 10.1016/j.smrv.2018.08.004

Takami, H., Okamoto, Y., Yamashita, H., Okada, G., & Yamawaki, S. (2007). Attenuated anterior cingulate activation during a verbal fluency task in elderly patients with a history of Multiple-episode depression. American Journal of Geriatric Psychiatry, 15(7), 594–603. doi: 10.1097/01.JGP.0b013e31802ea919

Tedeschini, E., Levkovitz, Y., Iovieno, N., Ameral, V. E., Nelson, J. C., & Papakostas, G. I. (2011). Efficacy of antidepressants for late-life depression: A meta-analysis and meta-regression of placebo-controlled randomized trials. The Journal of Clinical Psychiatry, 72(12), 1660–1668. doi: 10.4088/JCP.10r06531

Tench, C. R., Tanasescu, R., Cottam, W. J., Constantinescu, C. S., & Auer, D. P. (2019). Coordinate based meta-analysis does not show grey matter atrophy in narcolepsy. Neuroscience & Biobehavioral Reviews, 102, 427–429. doi: 10.1016/j.neubiorev.2016.10.019

Thompson, P. M., Jahanshad, N., Ching, C. R. K., Salminen, L. E., Thomopoulos, S. I., Bright, J., … Zelman, V. (2020). ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. Translational Psychiatry, 10(1), 1–28. doi: 10.1038/s41398-020-0705-1

Tittmann, M., Gunther, T., Sacher, J., Himmerich, H., Villringer, A., Hegerl, U., & Schonknecht, P. (2014). Structural Brain Changes in Early-Onset and Late-Onset Depression: An Update of Volumetric MRI Findings. International Journal of Imaging Systems and Technology, 24(2), 149–160. doi: 10.1002/ima.22089

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses. Human Brain Mapping, 33(1), 1–13. doi: 10.1002/hbm.21186

Vaishnavi, S., & Taylor, W. D. (2006). Neuroimaging in late-life depression. International Review of Psychiatry, 18(5), 443–451. doi: 10.1080/09540260600935454

van Agtmaal, M. J. M., Houben, A. J. H. M., Pouwer, F., Stehouwer, C. D. A., & Schram, M. T. (2017). Association of Microvascular Dysfunction With Late-Life Depression: A Systematic Review and Meta-analysis. JAMA Psychiatry, 74(7), 729–739. doi: 10.1001/jamapsychiatry.2017.0984

Vanyukov, P. M., Szanto, K., Siegle, G. J., Hallquist, M. N., Reynolds, C. F., Aizenstein, H. J., & Dombrovski, A. Y. (2015). Impulsive traits and unplanned suicide attempts predict exaggerated prefrontal response to angry faces in the elderly. American Journal of Geriatric Psychiatry, 23(8), 829–839. doi: 10.1016/j.jagp.2014.10.004

Wang, L., Krishnan, K. R., Steffens, D. C., Potter, G. G., Dolcos, F., & McCarthy, G. (2008). Depressive state- and disease-related alterations in neural responses to affective and executive challenges in geriatric depression. American Journal of Psychiatry, 165(7), 863–871. doi: 10.1176/appi.ajp.2008.07101590

Wang, Li, Leonards, C. O., Sterzer, P., & Ebinger, M. (2014). White matter lesions and depression: A systematic review and meta-analysis. Journal of Psychiatric Research, 56, 56–64. doi: 10.1016/j.jpsychires.2014.05.005

Weber, K., Giannakopoulos, P., Delaloye, C., de Bilbao, F., Moy, G., Ebbing, K., … Canuto, A. (2012). Personality traits, cognition and volumetric MRI changes in elderly patients with early-onset depression: A 2-year follow-up study. Psychiatry Research, 198(1), 47–52. doi: 10.1016/j.psychres.2011.11.017

Weber, K., Giannakopoulos, P., Delaloye, C., de Bilbao, F., Moy, G., Moussa, A., … Canuto, A. (2010). Volumetric MRI changes, cognition and personality traits in old age depression. Journal of Affective Disorders, 124(3), 275–282. doi: 10.1016/j.jad.2009.11.016

Wei, J., Hou, R., Zhang, X., Xu, H., Xie, L., Chandrasekar, E. K., … Goodman, M. (2019). The association of late-life depression with all-cause and cardiovascular mortality among community-dwelling older adults: Systematic review and meta-analysis. Br J Psychiatry, 215(2), 449–455. doi: 10.1192/bjp.2019.74

Weisenbach, S. L., Kassel, M. T., Rao, J., Weldon, A. L., Avery, E. T., Briceno, E. M., … Langenecker, S. A. (2014). Differential prefrontal and subcortical circuitry engagement during encoding of semantically related words in patients with late-life depression. International Journal of Geriatric Psychiatry, 29(11), 1104–1115. doi: 10.1002/gps.4165

Wen, M. C., Steffens, D. C., Chen, M. K., & Zainal, N. H. (2014). Diffusion tensor imaging studies in late-life depression: Systematic review and meta-analysis. International Journal of Geriatric Psychiatry, 29(12), 1173–1184. doi: 10.1002/gps.4129

Wong, N. M., Liu, H. L., Lin, C., Huang, C. M., Wai, Y. Y., Lee, S. H., & Lee, T. M. (2016). Loneliness in late-life depression: Structural and functional connectivity during affective processing. Psychological Medicine, 46(12), 2485–2499. doi: 10.1017/S0033291716001033

Wu, K.-Y., Hsiao, I.-T., Chen, C.-S., Chen, C.-H., Hsieh, C.-J., Wai, Y.-Y., … Lin, K.-J. (2014). Increased brain amyloid deposition in patients with a lifetime history of major depression: Evidenced on 18F-florbetapir (AV-45/Amyvid) positron emission tomography. European Journal of Nuclear Medicine and Molecular Imaging, 41(4), 714–722. doi: 10.1007/s00259-013-2627-0

Xie, C., Li, W., Chen, G., Douglas Ward, B., Franczak, M. B., Jones, J. L., … Goveas, J. S. (2012). The co-existence of geriatric depression and amnestic mild cognitive impairment detrimentally affect gray matter volumes: Voxel-based morphometry study. Behavioural Brain Research, 235(2), 244–250. doi: 10.1016/j.bbr.2012.08.007

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology, 106(3), 1125–1165. doi: 10.1152/jn.00338.2011

Yuan, Y., Zhang, Z., Bai, F., Yu, H., Shi, Y., Qian, Y., … Liu, Z. (2008). Abnormal neural activity in the patients with remitted geriatric depression: A resting-state functional magnetic resonance imaging study. Journal of Affective Disorders, 111(2–3), 145–152. doi: 10.1016/j.jad.2008.02.016

Yuan, Y., Zhu, W., Zhang, Z., Bai, F., Yu, H., Shi, Y., … Liu, Z. (2008). Regional Gray Matter Changes Are Associated with Cognitive Deficits in Remitted Geriatric Depression: An Optimized Voxel-Based Morphometry Study. Biological Psychiatry, 64(6), 541–544. doi: 10.1016/j.biopsych.2008.04.032

Yue, Y., Jia, X., Hou, Z., Zang, Y., & Yuan, Y. (2015). Frequency-dependent amplitude alterations of resting-state spontaneous fluctuations in late-onset depression. Biomed Research International, 2015. doi: 10.1155/2015/505479

Zang, Y., Jiang, T., Lu, Y., He, Y., & Tian, L. (2004). Regional homogeneity approach to fMRI data analysis. Neuroimage, 22(1), 394–400.

Zivin, K., Wharton, T., & Rostant, O. (2013). The economic, public health, and caregiver burden of late-life depression. Psychiatr Clin North Am, 36(4), 631–649. doi: 10.1016/j.psc.2013.08.008

Zou, Q.-H., Zhu, C.-Z., Yang, Y., Zuo, X.-N., Long, X.-Y., Cao, Q.-J., ... Zang, Y.-F. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. Journal of Neuroscience Methods, 172(1), 137–141.

Database	Search Query
PubMed	("voxel based morphometry" OR VBM OR "functional magnetic resonance imaging" OR fMRI OR "Positron-Emission Tomography" OR PET) AND (elderly[tiab] OR geriatric[tiab] OR "late life"[tiab] OR "later life"[tiab] OR "late onset"[tiab] OR older[tiab] OR "old age"[tiab]) AND (depress* OR MDD)
Embase	('voxel based morphometry'/exp OR 'voxel based morphometry' OR vbm OR 'functional magnetic resonance imaging'/exp OR 'functional magnetic resonance imaging' OR 'fmri'/exp OR fmri OR 'positron-emission tomography'/exp OR 'positron-emission tomography' OR 'pet'/exp OR pet) AND (elderly:ti,ab OR geriatric:ti,ab OR 'late life':ti,ab OR 'later life':ti,ab OR 'late onset':ti,ab OR older:ti,ab OR 'old age':ti,ab) AND (depress* OR mdd)
Scopusa	("voxel based morphometry" OR VBM OR "functional magnetic resonance imaging" OR fMRI OR "Positron-Emission Tomography" OR PET) AND (elderly OR geriatric OR "late life" OR "later life" OR "late onset" OR older OR "old age") AND (depress* OR MDD)
Web of Knowledgeb	("voxel based morphometry" OR VBM OR "functional magnetic resonance imaging" OR fMRI OR "Positron-Emission Tomography" OR PET) AND (elderly OR geriatric OR "late life" OR "later life" OR "late onset" OR older OR "old age") AND (depress* OR MDD)

Table S1: Search query used in each database.

[a] In title, abstract or keywords. [b] In 'topic'

	First Author (Year)a	Scanner Field Strength	Software	Multiple Comparison Correction	Covariates	Coordinates Space	Source of Coordinates
1	Albert, K (2017)	3T	SPM	AlphaSim (k=58, p <0.000001)	n.r.	MNI	Figure 3
2	Bobb, DS (2012)	3T	FSL	FWE	n.r.	MNI	Table 2
3	Bricenõ, EM (2015)	3T	FSL	AlphaSim (k=55, p<0.005)	n.r.	MNI	Table 3
	Rao, JA (2015)	3T	FSL, SPM	AlphaSim (k=264 mm3, p<0.003)	n.r.	MNI	Table 2
	Weisenbach, SL (2014)	3T	SPM	AlphaSim (k=264 mm3, p<0.003)	n.r.	MNI	Table 2
4	Byun, MS (2016)	3T	SPM	FWE (k=100)	age, gender, education level	MNI	Table S3
5	Chen, JD (2012)	1.5T	SPM	AlphaSim (k=675 mm3, p<0.005)	age, voxel-wise GM volume	MNI	Table 2
	Guo, W (2013)	1.5T	SPM	FDR	n.r.	MNI	Table 2
	Liu, F (2012)	1.5T	SPM	Not Corrected (k=1483 mm3, p <0.005)	age, voxel-wise GM volume	MNI	Table 2
6	De Asis, J (2001)	n.a.	SPM	Not Corrected (p <0.01)	global rCBF	Talairach	Text
7	Dombrovski, AY (2013)	3T	AFNI	AlphaSim (k=67)	n.r.	MNI	Author
8	Dumas, JA (2015)	3T	SPM	Not Corrected (k=200)	n.r.	MNI	Figure 1
9	Fang, J (2015)	1.5T	SPM	Corrected (method n.r.)	n.r.	MNI	Tables 2, 3
10	Harada, K (2016)	3T	SPM	FWE	n.r.	MNI	Table 2
	Harada, K (2018)	3T	SPM	FWE	n.r.	MNI	Text
11	Hou, Z (2016)	3T	SPM	AlphaSim (k=55, p <0.001)	GM volume, age, gender, education level	MNI	Table 2
12	Huang, CM (2019)	3T	SPM	FWE (p<0.001)	n.r.	MNI	Table 2
13	Hwang, JP (2010)	2T	SPM	Not Corrected (p <0.001)	age, education level, duration of illness	MNI	Table 2
14	Lee, TW (2013)	3T	AFNI	FWE (k=20, p<0.005)	n.r.	Talairach	Table 4
15	Li, J (2020)	3T	SPM	AlphaSim (p <0.001)	age, gender education level, MMSE score, framewise displacement	MNI	Table 2
16	Mah, L (2011)	3T	n.r.	n.r.	n.r.	Talairach	Text
17	Oudega, ML (2014)	1T	SPM	Not Corrected (k=50, p <0.001)	GM volume	MNI	Table 2
18	Respino, M (2019)	3T	SPM	FWE (p<0.001)	sex, education, mean framewise displacement	MNI	Table 2
19	Ribeiz, SRI (2013)	1.5T	SPM	FWE	GM volume, education level	Talairach	Text
20	Smith, GS (2009)	1.5T	SPM	Not Corrected (k = 50, p <001)	n.r.	MNI	Tables 1, 3b
21	Takami, H (2007)	n.r.	SPM	Not Corrected (p <0.001)	n.r.	Talairach	Table 3
22	Wang, L (2008)	4T	custom script	Not Corrected (k = 5, p <001)	age, education level, number of episodes, duration of disease	MNI	Table 2
23	Wong, NM (2016)	3T	FSL	FWE (p<0.001)	n.r.	MNI	Text
24	Xie, C (2012)	3T	SPM	FDR (k=150)	age, gender, education level, intracranial volume	MNI	Table 2
25	Yuan, Y (2008a)	1.5T	SPM, AFNI	AlphaSim (k=270 mm3, p<0.005)	n.r.	Talairach	Table 2
	Yuan, Y (2008b)	1.5T	SPM	Not Corrected (k=80 mm3 ,p<0.001)	n.r.	MNI	Table 2
26	Yue, Y (2015)	3T	SPM	Monte Carlo (k=2079 mm3, p <0.05)	n.r.	MNI	Table 2

Table S2: Technical details of studies included in the meta-analysis.

n.r. = not reported; n.a. = not applicable; T = Tesla; SPM = Statistical Parametric Mapping; FSL = FMRIB Software Library; AFNI = Analysis of Functional NeuroImages; FWE = family-wise error correction; FDR = false discovery rate; GM = grey matter; rCBF = relative cerebral blood flow; MMSE = mini-mental state examination; MNI = Montreal Neurological Institute; [a] Publications with overlapping samples are grouped together