Systematic review and meta-analyses of neural structural and functional differences in generalized anxiety disorder and healthy controls using magnetic resonance imaging

Tiffany A. Kolesar, Elena Bilevicius, Alyssia D. Wilson, Jennifer Kornelsen

Objective: To compare structure, functional connectivity (FC) and task-based neural differences in subjects with generalized anxiety disorder (GAD) compared to healthy controls (HC).

Methods: The Embase, Ovid Medline, PsycINFO, Scopus, and Web of Science databases were searched from inception until March 12, 2018. Two reviewers independently screened titles, abstracts, and full-text articles. Data were extracted from records directly contrasting GAD and HC that included structure (connectivity and local indices such as volume, etc.), FC, or task-based magnetic resonance imaging data. Meta-analyses were conducted, as applicable, using AES-SDM software.

Results: The literature search produced 4,645 total records, of which 85 met the inclusion criteria for the systematic review. Records included structural (n=35), FC (n=33), and task-based (n=42) findings. Meta-analyses were conducted on voxel-based morphometry and task-based results.

Discussion: The systematic review confirms and extends findings from previous reviews. Although few whole-brain resting state studies were conducted, key nodes of resting state networks have altered physiology: the hippocampus (default network), ACC and amygdala (salience network), have reduced volume, and the dlPFC (central executive network) and ACC have reduced FC with the amygdala in GAD. Nodes in the sensorimotor network are also altered with greater pre- and postcentral volume, reduced supplementary motor area volume, and reduced FC in anterior and increased FC in posterior cerebellum.

Conclusions: Despite limitations due to sample size, the meta-analyses highly agree with the systematic review and provide evidence of widely distributed neural differences in subjects with GAD, compared to HC. Further research optimized for meta-analyses would greatly improve large-scale comparisons.

1. Introduction

Anxiety disorders are highly prevalent in the general population, and generalized anxiety disorder (GAD) is one of the most common forms (Somers et al., 2006). GAD is characterized by chronic, persistent worry that is present more days than not over at least the past six months (American Psychiatric Association, 2013). In addition to the psychological manifestation of this disorder, GAD also presents physically. In fact, it is often physical ailments—such as gastrointestinal upset or headaches—that cause patients to seek treatment (Stein and Sareen, 2015). According to the Diagnostic and Statistical Manual for Mental Disorders, 5th edition (DSM-5), an adult patient’s chronic worry must be accompanied by three or more of the following symptoms—irritability, difficulty concentrating, insomnia, fatigue, restlessness, or muscle tension—again occurring more often than not in the past 6 months for a GAD diagnosis to be made (American Psychiatric Association, 2013). A comprehensive systematic review and meta-analysis of the body of work to date may elucidate the common neural correlates underlying this disorder. The purpose of the current work is to review the neural differences occurring in GAD, compared to healthy controls (HC), as assessed by structural and functional magnetic resonance imaging (fMRI) studies.

Neurophysiology can be assessed various ways, even within the field of magnetic resonance imaging (MRI). Investigations of brain structure...
commonly include measures of local volumetric (e.g., voxel-based morphometry), cortical thickness, and surface area differences and, less common, local gyriﬁcation index (i.e., cortex within sulcal folds, compared to gyral cortex) and white matter lesions (hyperintensities in a typical T2-weighted MRI). Furthermore, physical white matter connections can also be assessed—diffusion tensor imaging (DTI) assesses this structural connectivity via fractional anisotropy (a measure of sphericity of diffusion in neural tissue), mean diffusivity (average diffusion within a region), apparent diffusion coefﬁcient (magnitude of diffusion in a region), tractography (a technique for modelling neural tracts), and axial (difffusivity along the principal axis) and radial difffusivity (average diffusivity along two minor axes). In addition to investigating structural neuroanatomy, much MRI research has been done elucidating neural function via task-based activation and functional connectivity (FC). Task-based fMRI identiﬁes regions of the brain or spinal cord whose activity correlates with task performance. FC assesses how the activity of various regions correlate to each other (Friston, 2011). Various measures of FC exist: Psychophysiological interaction (PPI) examines interactions between physiological variables and experimental (e.g., task) factors (Friston, 2011), regional homogeneity (ReHo) investigates local FC, evaluating the time-series of voxels and their nearest neighbours (Zang et al., 2004), amplitude of low frequency ﬂuctuations (ALFF) examines differences in the magnitude of the slow oscillating activity observed in resting state fMRI between regions, and individuals (Zang et al., 2007) and independent component analysis (ICA) identiﬁes signals with maximum independence from each other and can be used to separate resting state networks from each other (i.e., resting state fMRI; Calhoun et al., 2009).

Several reviews have been conducted in attempts to amalgamate results from the types of neuroimaging studies described above, in order to visualise how anxious brains differ from non-anxious ones. Recent reviews indicate that anxiety and mood disorders often share a common neurobiological pathophysiology involving the prefrontal cortex (PFC), hippocampus, and amygdala (Duval et al., 2015), with a key feature being increased amygdala and decreased PFC activity (Quide et al., 2012). In one review, fear-based conditions (panic disorder [PD]/speciﬁc phobias) resulted in greater involvement in emotion-generating regions (e.g., dorsal anterior cingulate cortex [ACC], amygdala, insula), while anxiety-based conditions (GAD/posttraumatic stress disorder [PTSD]) had greater PFC dysregulation (Duval et al., 2015).

Looking speciﬁcally at GAD, altered function was observed in the PFC and ACC resulting from tasks investigating emotion dysregulation, conditioned fear overgeneralization, and worry induction in one systematic review (Mochcovitch et al., 2014). Furthermore, reduced FC between the amygdala and cortex was also reported (Mochcovitch et al., 2014). Similarly, Hilbert et al. (2014), reviewing many of the same papers, observed alterations in the same three areas in GAD (PFC, amygdala, ACC), with the addition of the hippocampus. The main ﬁndings from Hilbert and colleagues’ systematic review were that GAD patients had abnormal activity in PFC and amygdala, increased amygdala grey matter (GM), and decreased FC and structural connectivity between these regions, combined with increased reactivity of the noradrenergic system, compared to HC. More recently, Fonzo and Etkin (2017) also observed abnormal PFC and limbic activation in response to facial affect processing, affective learning and regulation, and perseverative cognition tasks and altered FC when comparing GAD and HC groups. Although these results appear vague and nondescript (i.e., “abnormal” activity rather than increased or decreased), Fonzo and Etkin (2017) discussed that this variability may actually be a facet of “abnormal” activity rather than increased or decreased, Fonzo and Etkin (2017) discussed that this variability may actually be a facet of variability. The authors discuss that, because the pathological worry in GAD can be generated without external stimulation, this neural state may remain less impacted by external stimuli. All three of these systematic reviews come to the same conclusion: (f)MRI provides evidence for top-down emotion processing deﬁcits in GAD. Since these reviews were conducted (Fonzo and Etkin 2017; Hilbert et al. 2014; Mochcovitch et al., 2014), a large number of new studies have been published. Furthermore, no current papers have conducted meta-analyses on any aspect of GAD MRI work.

The purpose of the current systematic review and meta-analyses is to summarize all MRI studies that compare neural differences between subjects with GAD and HC, yielding structural, FC, or task-based results. We hypothesize that the results from the meta-analysis and systematic review will corroborate the ﬁndings of the previous systematic reviews conducted with fewer records, as well as identify regions previously under-recognized. The outcomes of this paper will be structural (local and connectivity measures), FC, and task-based activity from (f) MRI research in GAD and HC. The resulting synthesis will provide a more detailed understanding of the neurophysiology underlying this highly prevalent and debilitating anxiety disorder.

2. Methods

2.1. Literature search and selection criteria

The GAD neuroimaging literature was systematically searched on March 12, 2018, from inception. The comprehensive search included Medical Subject Headings, text, and keywords using the Embase, Ovid Medline, PsycINFO, Scopus, and Web of Science databases. Two main themes were included in the search: (1) MRI and (2) generalized anxiety disorder (please see supplemental material for the full search terms). Note that different search terms were used for different databases, based on the requirements of each database—for example, databases that use Medical Subject Headings have specific terminology that may not be applicable to other databases. The reference lists of all included articles were reviewed to identify further relevant papers. Studies were included if they were full-text, published articles that reported on original research using MRI with human subjects and if they compared neural structure (connectivity and local indices—e.g., volume), FC, or activity in subjects with GAD to HC. Although country of origin was not restricted, language was restricted to English.

2.2. Study selection

All titles and abstracts were reviewed independently by two reviewers (T.A.K. and E.B.) using EndNote X7 software. Any title or abstract selected by either reviewer was included for further examination. All full-text articles were then screened for final inclusion by the same two reviewers; any disagreements at this stage were solved by consensus. Full-text articles were included for final selection if they met the following criteria: (1) original research; (2) not solely an abstract; (3) reported human MRI ﬁndings; (4) in a GAD population where GAD was the primary or most prominent diagnosis; (5) included a contrast between GAD and HC participants. While the systematic review portion of the current work includes whole-brain, region-of-interest, and seed-based results, the meta-analyses are limited to studies that included whole-brain data.

2.3. Data extraction and synthesis

Data were extracted using a standardized form, including the publication year, sample size, populations sampled (some studies included additional diagnoses), study modality (structure, FC, task), comorbidities, disease duration, diagnostic criteria, medications, questionnaires, MRI sequence type, data analysis software, contrasts performed, and regions (including coordinates, Brodmann areas, and lateralisation, as applicable) of structural, FC, and activity differences (see supplementary data spreadsheet). Demographic data included distribution of sex, handedness, age, and location of data collection. Attempts were made to contact authors to obtain missing information; however, if authors could not be reached, information remains incomplete in some instances.
2.4. Meta-Analyses

Two meta-analyses were conducted: one for voxel-based morphometry (VBM), and one for task-based results (comparing neutral and negative emotion-evoking stimuli) using Anisotropic Effect Size Seed-Based D Mapping (AES-SDM) software, version 5.15 (www.sdmproject.com; Radua and Mataix-Cols, 2012; Radua et al., 2012, 2014). Instead of assigning voxels a conventional value, this software uses Hedge’s g to assign each voxel a measure of effect size (Radua et al., 2012). This software has been used to assess a variety of structural and functional MRI findings from various populations in the past (e.g., Jiang et al., 2017; Pico-Perez et al., 2017; Wang et al., 2018). Records were included in meta-analyses only if they explored the whole brain, and used a single significance threshold throughout the brain (Radua and Mataix-Cols, 2012). Additionally, if multiple studies were individually eligible for meta-analysis, but had confirmed or suspected participant overlap, the record with a greater sample size was included in the meta-analysis. When possible, whole brain maps were used, while peak voxels were used when maps were not available. Furthermore, our criterion for meta-analysis was a minimum of 5 studies, provided they included at least one whole-brain map. Although some records included results with a patient group in addition to GAD and HC and were eligible for the systematic review, in some cases it was not possible to isolate results specific to only GAD and HC groups, these records were excluded from the meta-analysis (e.g., Ball et al., 2013; Blair et al., 2012; Fonzo et al., 2015). Studies were included in the meta-analysis if they reported null findings, if they met the eligibility criteria.

First, meta-analyses that included whole-brain maps were converted to a usable format for the AES-SDM software. In one task-based study (Palm et al., 2011), three contrasts were performed comparing negative emotion-evoking faces to a neutral baseline (fearful > neutral, angry > neutral, sad > neutral). As it would not be appropriate to add these contrasts to the meta-analysis as individual records—this would bias the results by including data from the same individuals as if they were independent—the peak coordinates from these three contrasts were combined into a single brain map so that all of the data from these negative contrasts could be used in the meta-analysis. This combined brain map was then preprocessed along with the remaining task records. For both meta-analyses, any values listed as z-scores were converted to t-scores prior to preprocessing. Data from each meta-analysis was preprocessed using 50 Monte Carlo randomizations. Next, a voxel-wise random-effects analysis was conducted in which the weighted mean differences in GM or activity between subjects with GAD and HC were computed, providing between-study heterogeneity estimates, variance (\(I^2\)), \(z\), and probability maps. This mean analysis is weighted for sample size, intra-study variance, and between group heterogeneity (Radua and Mataix-Cols, 2009, 2012; Radua et al., 2014). Due to the low sample sizes of the meta-analyses, complementary meta-analyses were limited to jackknife sensitivity analyses, as such analyses looking at age-, medication-, or comorbidity-effects were not conducted. Statistical significance was set to \(p_{voxel} (< 0.005, uncorrected)\), with peak SDM-\(z\) score > 1, and a minimum extent of 10 contiguous voxels, for optimal balance between \(\alpha\) and \(\beta\) errors (Radua and Mataix-Cols, 2012).

2.5. Assessment of study consistency

Consistency was assessed qualitatively for the systematic review. The included studies varied in a number of areas, particularly in inclusion/exclusion criteria as various age groups, comorbidities, medication use, and diagnostic criteria were either allowed or disallowed. Additionally, study design was highly varied across studies, which is not unexpected, particularly amongst task-based studies.

Upon examination of the systematic review data, many of the cerebellum results were simply labelled as ‘cerebellum’ and more detailed descriptions were not provided, perhaps attributable to software limitations. To develop a better understanding of cerebellar location, all
cerebellum coordinates were labelled using either Talairach Client (for Talairach coordinates; http://www.talairach.org/client.html) or the aal atlas in MRICron (for Montreal Neurological Institute [MNI] coordinates; https://www.nitrc.org/projects/mricron).

For the meta-analyses, robustness of findings was assessed using jackknife sensitivity analyses which use a leave-one-out method (Radua and Mataix-Cols, 2009). I² index and Egger's tests, used to assess heterogeneity of effect sizes and publication bias, respectively, were also conducted for each meta-analysis. Funnel plots were created for significant meta-analytic clusters.

3. Results

3.1. Identification of studies

The search strategy yielded 4645 total records, and after 1206 duplicates were removed, 85 met the inclusion criteria (see Fig. 1 for flow diagram). Of the included records, 35 included structural analyses (Abdallah et al., 2013; Andreescu et al., 2017; Brambilla et al., 2012; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016b; Chen and Etkin, 2013; De Bells et al., 2000, 2002; Etkin et al., 2009; Hettema et al., 2012; Hilbert et al., 2015; Karim et al., 2016; Liao et al., 2013, 2014a, 2014b; Makovac et al., 2016a; Mohlman et al., 2009; Molen et al., 2017; Moon and Jeong, 2016a, 2016b, 2017a, 2017b; Moon et al., 2015a, 2015b; Mueller et al., 2013; Schienle et al., 2011; Strawn et al., 2014, 2013; Terlevec et al., 2013; Tromp et al., 2012; Wang et al., 2016b; Zhang et al., 2011; Zhang et al., 2013), 32 included FC analyses (Andreescu et al., 2015; Andreescu et al., 2014; Buff et al., 2016; Cha et al., 2014a; Cha et al., 2016b; Chen and Etkin, 2013; Cui et al., 2014, 2016; Etkin et al., 2016; Etkin et al., 2009; Etkin and Schatzberg, 2011; Fonzo et al., 2014; Greenberg et al., 2013; Hölzle et al., 2013; Laufer et al., 2016; Li et al., 2018; Li et al., 2016; Liu et al., 2015; Makovac et al., 2016b; Makovac et al., 2018; McClure et al., 2007; Mohlman et al., 2017; Monk et al., 2008; Oathes et al., 2015; Pace-Schott et al., 2017; Qiao et al., 2017; Rabany et al., 2017; Roy et al., 2013; Strawn et al., 2012; Toazza et al., 2016; Tromp et al., 2012; Wang et al., 2016a; Xia et al., 2017), and 42 included task-based designs (Andreescu et al., 2011, 2015; Ball et al., 2013; Blair et al., 2008; Blair et al., 2012; Blair et al., 2017; Buff et al., 2016; Carlson et al., 2017; Cha et al., 2014a; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016b; Chen and Etkin, 2013; Diwadkar et al., 2017; Etkin et al., 2010; Etkin and Schatzberg, 2011; Fitzgerald et al., 2017; Fonzo et al., 2015, 2014; Greenberg et al., 2013; Guyer et al., 2012; Hölzle et al., 2013; Karim et al., 2016; Laufer et al., 2016; Makovac et al., 2018; McClure et al., 2007; Mohlman et al., 2017; Monk et al., 2006, 2008; Moon and Jeong, 2015b, 2017b; Moon et al., 2016, 2015b, 2017; Nitschke et al., 2009; Ottaviani et al., 2016; Palm et al., 2011; Paulesu et al., 2010; Price et al., 2011; Straw et al., 2012; Whalen et al., 2008; White et al., 2017; Vassa et al., 2012). For reader ease, records are sorted by modality in the supplementary data spreadsheet. For information on any specific study, refer to the supplementary data spreadsheet.

3.2. Details of included studies

Although databases were searched from inception, studies in which GAD was investigated with MRI ranged from 2008 to 2018. Out of the 85 records included, 1 was conducted in South America, 14 were conducted in Europe, 23 in Asia, and 47 in North America (see Table 1 for references). Handenedness was recorded in 43 of the papers (see Table 1), of which 99% of the participants were right-handed. Across the 85 studies, there were a total of 4160 participants (1855 with a diagnosis of GAD) that underwent an MRI scan with approximately 63% of participants being female. However, this sample size is inflated as many papers shared participants within labs (see supplementary Table S1).

Sixteen studies included more than one patient population (i.e., in addition to a GAD group), including social anxiety disorder (SAD; also including previous iterations such as generalized social phobia and social phobia), PD, major depressive disorder (MDD), PTSD, and primary insomnia (see Table 1). Additionally, of these 16 studies, five included a purposeful comorbid group in which patients had both GAD, and generalized social phobia or MDD comorbidity. These studies included these comorbidities or differential diagnoses as distinct groups, rather than simply allowing comorbidities in the inclusion criteria; i.e., many of the included studies did not exclude participants for having additional anxiety disorders or mood disorders. Two records compared anxiety disorders in general to HC, but were included as they conducted contrasts with the GAD subgroup in their anxiety group (Mueller et al., 2013; Toazza et al., 2016). For additional information on comorbidities, see the supplementary data spreadsheet.

All records included mean, median or range of participant ages: 16 studies were done in an adolescent population, 61 were done in an adult population, 7 studies were done in an elderly population, and adult and elderly participants were compared in 1 study (see Table 1).

3.3. Study design

Structural analyses were conducted in 35 records and spanned a variety of methodologies, including (1) diffusion tensor imaging (DTI, \( n = 10 \)), (2) white matter hyperintensity (WMH, \( n = 2 \)), (3) cortical thickness analysis (CTA, \( n = 4 \)), (4) VBM \( (n = 16) \), (5) other volumetric analyses \( (n = 10) \), (6) surface area \( (n = 1) \), and (7) local gyrification index \( (n = 1) \); see Table 2). FC analyses were conducted in 33 records: resting state fMRI scans were used in 12—defined here as a separate fMRI scan, acquired in the absence of a task, using basic seed-based, region-of-interest or independent components analyses (ICA). Six studies included measures of FC conducted from task-based data and 10 studies included psychophysiological interaction (PPI; 2 observed no significant results Cha et al., 2016b; Greenberg et al., 2013), however, between-groups contrasts were not conducted for PPI in one record (Laufer et al., 2016). A few records included FC analyses for hierarchical partner matching-ICA \( (n = 1) \), amplitude of low frequency fluctuations analyses \( (n = 1) \), effective connectivity \( (n = 2) \), and regional homogeneity \( (n = 2) \); see Table 2). Finally, 42 records included a task, and these were separated into groups including: (1) null judgement/passive (discerning characteristics of no interest to the researchers like gender or nose width, or simply viewing emotional stimuli), (2) congruency and conflict (deciphering congruent and incongruent stimuli), (3) emotion modulation (maintaining or altering emotions after neutral or anxiety-inducing distractors)), and (6) miscellaneous tasks (see Table 2). For more specific task information, please see the supplementary data spreadsheet. In one record, two distinct tasks were performed (Blair et al., 2012), and these are listed separately in Table 2. To focus the review, neuroimaging results obtained from correlation with questionnaires or behavioural data are not reported here. For this reason, results are omitted from 2 records as the only significant results were found after co-varying neural activity with questionnaire data (Karim et al., 2016; Mohlman et al., 2009).

3.4. Systematic review results

Common MRI results for comparisons between subjects with GAD and HC can be found in supplementary Table S2. Regions were listed in Table S2 if they were found in at least two records from different laboratories, but a full list of results can be found in the supplementary data spreadsheet. The most commonly occurring regions include the same four regions consistently identified by other systematic reviews: the dorsolateral PFC (dlPFC), ACC, amygdala, and hippocampus.

The results from the ACC were largely mixed: results indicate both increased \( (n = 6) \); Andreescu et al., 2011; Fonzo et al., 2014; Laufer.
| Sample Characteristic | N | Records |
|-----------------------|---|---------|
| **Location**          |   |         |
| South America         | 1 | (Toazza et al., 2016) |
| Europe                | 14 | (Brambilla et al., 2012; Buff et al., 2016; Diwadkar et al., 2017; Hilbert et al., 2015; Laufer et al., 2016; Makovac et al., 2016a, 2016b; Mohlman et al., 2013; Molent et al., 2013; Nitschke et al., 2009; Ottaviani et al., 2015; Palm et al., 2011; Paulesu et al., 2010; Schielen et al., 2011; Terlevic et al., 2013) |
| Asia                  | 23 | (Cui et al., 2016; Li et al., 2018; Li et al., 2016; Liao et al., 2015, 2015a, 2015b; Liu et al., 2015; Moon et al., 2015a; Moon et al., 2014, 2015a; Moon et al., 2015b; Moon et al., 2017; Qiao et al., 2017; Wang et al., 2016; Wang et al., 2016b; Xin et al., 2017; Zhang et al., 2011, 2013) |
| North America         | 47 | (Abdalalah et al., 2013; Andreescu et al., 2011, 2015, 2014; Andreescu et al., 2017; Ball et al., 2013; Blair et al., 2008; Blair et al., 2012; Blair et al., 2017; Carlson et al., 2017; Cha et al., 2014a; Cha et al., 2014b; Cha et al., 2016a; Cha et al., 2016b; Chen and Etkin, 2013; De Bellis et al., 2000, 2002; Etkin et al., 2009, 2012; Etkin and Schatzberg, 2011; Fitzgeral et al., 2017; Fonzo et al., 2013; Fonzo et al., 2015a; Fonzo et al., 2015b; Fonzo et al., 2016; Fonzo et al., 2017; Fous et al., 2015; Gambling et al., 2013; Guyer et al., 2012; Hettema et al., 2012; Holzman et al., 2013; Karim et al., 2016; McClure et al., 2007; Mohlman et al., 2017, 2009; Monk et al., 2006, 2008; Mueller et al., 2013; Nitschke et al., 2009; Oathes et al., 2015; Pace-Schott et al., 2017; Price et al., 2011; Rabany et al., 2017; Roy et al., 2013; Strawn et al., 2012, 2014; Tromp et al., 2012; Whalen et al., 2008, 2017; Yassa et al., 2012) |
| **Handedness**        |   |         |
| Recorded              | 43 | (Brambilla et al., 2012; Carlson et al., 2017; Cui et al., 2016; De Bellis et al., 2002; Diwadkar et al., 2017; Etkin et al., 2010, 2009; Etkin and Schatzberg, 2011; Hettema et al., 2012; Hilbert et al., 2015; Holzel et al., 2013; Li et al., 2018; Li et al., 2016; Liao et al., 2013, 2014a, 2014b; Liu et al., 2015; Makovac et al., 2016a, 2016b; Makovac et al., 2018; Mohlman et al., 2017, 2009; Monk et al., 2008; Moon et al., 2015a; Moon et al., 2015b; Moon et al., 2017; Nitschke et al., 2009; Ottaviani et al., 2016; Pace-Schott et al., 2017; Paulus et al., 2010; Price et al., 2011; Qiao et al., 2017; Roy et al., 2013; Schielen et al., 2011; Terlevic et al., 2013; Toazza et al., 2016; Tromp et al., 2012; Wang et al., 2016a; Wang et al., 2016b; Whalen et al., 2008; Yassa et al., 2012; Zhang et al., 2011, 2013) |
| **Sample Age**        |   |         |
| Adolescent            | 16 | (De Bellis et al., 2000, 2002; Guyer et al., 2012; Liao et al., 2013, 2014a, 2014b; Liu et al., 2015; McClure et al., 2007; Monk et al., 2006, 2008; Mueller et al., 2013; Roy et al., 2013; Strawn et al., 2012, 2014, 2013; Toazza et al., 2016) |
| Adult                 | 61 | (Abdalalah et al., 2013; Ball et al., 2013; Blair et al., 2008; Blair et al., 2012; Blair et al., 2017; Brambilla et al., 2012; Buff et al., 2016; Carlson et al., 2017; Cha et al., 2014a; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016c; Chen and Etkin, 2013; Cui et al., 2016; Diwadkar et al., 2017; Etkin et al., 2016, 2009; Etkin and Schatzberg, 2011; Fitzgeral et al., 2017; Fonzo et al., 2015a, 2015b; Greenberg et al., 2013, 2009; Hettema et al., 2012; Hilbert et al., 2015; Holzel et al., 2013; Lauffer et al., 2016b; Li et al., 2018; Li et al., 2016; Makovac et al., 2016c, 2016b; Makovac et al., 2018; Molent et al., 2013, 2017; Moon and Jeong, 2015a, 2015b, 2017a, 2017b; Moon et al., 2013, 2015a, 2015b, 2016, 2016a, 2016b; Moon et al., 2015a; Moon et al., 2016b; Whalen et al., 2008; White et al., 2017; Xin et al., 2017; Yassa et al., 2012; Zhang et al., 2011, 2013) |
| Elderly               | 7  | (Andreescu et al., 2011, 2015, 2017; Karim et al., 2016; Mohlman et al., 2017, 2009; Price et al., 2011) |
| Adult + Elderly       | 1  | (Andreescu et al., 2014) |

**Records with Additional Patient Groups**

| Group                  | N | Records |
|------------------------|---|---------|
| GSP                    | 1 | (Blair et al., 2008) |
| GSP + GAD/GSP          | 1 | (Blair et al., 2012) |
| SP                     | 1 | (Guyer et al., 2012) |
| SAD                    | 2 | (Blair et al., 2017; Rabany et al., 2017) |
| SAD + PD               | 2 | (Buff et al., 2016; Fonzo et al., 2015) |
| PD                     | 3 | (Ball et al., 2013; Cui et al., 2016; Terlevic et al., 2013) |
| GAD/MDD                | 1 | (Cha et al., 2016a) |
| GAD/MDD + MDD          | 3 | (Carlson et al., 2017; Etkin, Schatzberg, 2011; Oathes et al., 2015) |
| PTSD                   | 2 | (Chen and Etkin, 2013; Zhang et al., 2011) |
| Primary Insomnia       | 1 | (Pace-Schott et al., 2017) |

A '+' symbol indicates multiple patient groups, while a '/' indicates comorbid groups. Adolescent = ages 11–18; Adult = ages 19–59; Elderly = ages 60+; GSP = generalized social phobia; GAD = generalized anxiety disorder; SP = social phobia; SAD = social anxiety disorder; PD = panic disorder; MDD = major depressive disorder; PTSD = posttraumatic stress disorder.
Table 2
Study design and task-based stimuli used in included records.

| Modality                   | N   | References                                                                 |
|----------------------------|-----|-----------------------------------------------------------------------------|
| **Structure**              | 35  |                                                                             |
| Diffusion Tensor Imaging   | 10  | (Brambilla et al., 2012; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016b; Hettema et al., 2012; Liao et al., 2014b; Trompet al., 2012; Wang et al., 2016b; Zhang et al., 2011; Zhang et al., 2013) |
| Fractional Anisotropy      | 8   | (Cha et al., 2016b; Cha et al., 2014b; Hettema et al., 2012; Liao et al., 2014b; Trompet al., 2012; Wang et al., 2016b; Zhang et al., 2011; Zhang et al., 2013) |
| Mean Diffusivity           | 3   | (Cha et al., 2016a; Trompet al., 2012; Wang et al., 2016b)                  |
| Apparent Diffusion Coefficient | 1  | (Brambilla et al., 2012)                                                  |
| Tractography               | 1   | (Wang et al., 2016b)                                                      |
| White Matter Hyperintensity| 2   | (Andreescu et al., 2017; Karim et al., 2016)                               |
| Cortical Thickness Analysis| 4   | (Andreescu et al., 2017; Cha et al., 2014b; Molent et al., 2017; Strawn et al., 2014) |
| Voxel-Based Morphometry    | 16  | (Chen and Etkin, 2013; Etkin et al., 2009; Hilbert et al., 2015; Liao et al., 2013, 2014a; Makovac et al., 2016a; Moon and Jeong, 2015a, 2016, 2017a, 2017b; Moon et al., 2015a; Moon et al., 2014, 2015b; Mueller et al., 2013; Schienle et al., 2011; Strawn et al., 2013) |
| Volume                     | 10  | (Abdallah et al., 2013; Andreescu et al., 2017; Cha et al., 2014b; De Bels et al., 2000, 2002; Hettema et al., 2012; Karim et al., 2016; Mohlman et al., 2009; Molent et al., 2017; Terleer et al., 2013) |
| Surface Area               | 1   | (Molent et al., 2017)                                                     |
| Local Gyrification Index   | 1   | (Molent et al., 2017)                                                     |
| Functional Connectivity    | 33  |                                                                             |
| Resting State              | 12  | (Andreescu et al., 2014; Chen and Etkin, 2013; Cui et al., 2016; Etkin et al., 2009; Li et al., 2016; Liu et al., 2015; Oathes et al., 2015; Pace-Schott et al., 2017; Rabany et al., 2017; Roy et al., 2013; Toazza et al., 2016; Wang et al., 2016a) |
| Task-Related FC            | 7   | (Andreescu et al., 2015; Cha et al., 2014b; Hözel et al., 2013; Makovac et al., 2016b; Makovac et al., 2018; McClure et al., 2007; Strawn et al., 2012) |
| Psychophysiological Interaction | 10 | (Qiao et al., 2017)                                                      |
| HIF                | 1   | (Wang et al., 2016a)                                                    |
| Effective Connectivity    | 2   | (Mohlman et al., 2017; Qiao et al., 2017)                                 |
| Regional Homogeneity       | 2   | (Li et al., 2018; Xia et al., 2017)                                       |
| Task                        | 42  |                                                                             |
| Null Judgement/ Passive    | 10  | (Blair et al., 2008; Buff et al., 2016; Carlsson et al., 2017; Chen and Etkin, 2013; Etkin and Schatzberg, 2011; Fitzgibbon et al., 2017; Hözel et al., 2013; McClure et al., 2007; Nitschke et al., 2009; Palm et al., 2011; Whalen et al., 2008) |
| Passive view ‘Lost’ episode |    | (Carlsson et al., 2017)                                                 |
| Passively view cued emotional images |    | (Nitschke et al., 2009)                                                |
| Passively view emotional faces |    | (Whalen et al., 2008)                                                  |
| Passively view or appraise IAPS affect |    | (Fitzgibbon et al., 2017)                                             |
| Appraise face affect       |    | (Hözel et al., 2013)                                                   |
| Face hostility/nose width judgements |    | (McClure et al., 2007)                                               |
| Emotional faces + gender judgement |    | (Blair et al., 2008; Palms et al., 2013)                                |
| Emotional faces + face colour judgement |    | (Chen and Etkin, 2013)                                                |
| IAPS + blurry image judgement |    | (Buff et al., 2016)                                                   |
| Congruency/ Conflict       | 9   | (Blair et al., 2012; Etkin et al., 2010; Etkin and Schatzberg, 2011; Fitzgibbon et al., 2017; Hözel et al., 2013; McClure et al., 2007; Nitschke et al., 2009; Palm et al., 2011; Whalen et al., 2008) |
| Top-down attention control |    | (Blair et al., 2012)                                                   |
| Emotional conflict task    |    | (Etkin et al., 2010; Etkin and Schatzberg, 2011) | |
| Emotion face assessment task |    | (Fonzo et al., 2014)                                                  |
| Modified emotion assessment task |    | (Fonzo et al., 2013)                                                  |
| Congruent emotional faces  |    | (Karim et al., 2016; Monk et al., 2006, 2008) | |
| Emotional Stroop task      |    | (Price et al., 2011)                                                   |
| Emotion Modulation         | 6   | (Andreescu et al., 2011, 2015; Ball et al., 2015; Blair et al., 2012; Mohlman et al., 2017; Paulesu et al., 2010) |
| Worry induction or suppression |    | (Andreescu et al., 2011)                                             |
| Worry induction or neutral  |    | (Mohlman et al., 2017; Paulesu et al., 2010) | |
| Worry induction            |    | (Andreescu et al., 2015)                                              |
| Maintain or reduce reactions to images |    | (Ball et al., 2013)                                                   |
| Explicit emotion regulation |    | (Blair et al., 2012)                                                  |
| Conditioned Fear           | 6   | (Cha et al., 2014a; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016b; Greenberg et al., 2013; Laufer et al., 2016) |
| Fear generalisation (shape + shock) |    | (Cha et al., 2014a; Cha et al., 2016a; Cha et al., 2014b; Cha et al., 2016b; Greenberg et al., 2013) |
| Fear generalisation (tones + money) |    | (Laufer et al., 2016)                                                  |
| Memory                     | 6   | (Ovisalikar et al., 2017; Moon and Jeong, 2015b, 2017b; Moon et al., 2015b, 2017b) |

(continued on next page)
Finally, results indicated that subjects with GAD had reduced dlPFC volume (n = 5; Andreescu et al., 2017; Moon and Jeong, 2015a, 2016, 2017a, 2017b).

The results for the amygdala were somewhat clearer: all structural studies consistently showed increased volume (De Bellis et al., 2000; Etkin et al., 2009; Schienle et al., 2011) and FA (Zang et al., 2012) for subjects with GAD. While one study showed reduced effective connectivity in the amygdala (Yassa et al., 2012), another observed reduced FC between the right and left amygdala (Liu et al., 2015), and other FC results were greater for GAD (albeit with inconsistent seed regions; Andreescu et al., 2015; Buff et al., 2016; Liu et al., 2015; Mohlman et al., 2017; Qiao et al., 2017) and spanning all age groups. Finally, the majority of task results (n = 11) indicated greater amygdala activity for subjects with GAD for passive (Fitzgerald et al., 2017; Hözel et al., 2013; McClure et al., 2007; Nitschke et al., 2009), congruency (Etkin et al., 2010; Etkin and Schatzberg, 2011; Fonzo et al., 2015), and decreased activity in Ball et al., 2013; and greater activity in Mohlman et al., 2017; Price et al., 2011), and emotion modulation (Mohlman et al., 2017) tasks, while only a few studies in adults (n = 2) showed reduced activity for subjects with GAD in passive (Carlson et al., 2017) and congruency (Blair et al., 2012) tasks. One study investigating high uncertainty observed both increased and decreased activity in the amygdala (Yassa et al., 2012). Although these amygdala results included expected responses to aversive stimuli, it also included results for neutral stimuli in two cases (Hözel et al., 2013; Nitschke et al., 2009). Additionally, a study of regions that hypothesized amygdala volume (Hettema et al., 2012; Liao et al., 2013; Makovac et al., 2016a; Mohlman et al., 2009; Mueller et al., 2013) activity (Chen and Etkin, 2013; Whalen et al., 2008), or FC (Cha et al., 2016b; Greenberg et al., 2013; Laufer et al., 2016; Rabany et al., 2017) differences did not observe them. Finally, the hippocampus results were left-lateralized (with exceptions in: Abdallah et al., 2013 (bilateral); Cha et al., 2014a; Wang et al., 2016a) and indicated that subjects with GAD had reduced volume (Abdallah et al., 2013; Hettema et al., 2012; Moon and Jeong, 2017a; Moon et al., 2014; Moon et al., 2015b) and increased mean diffusivity (Cha et al., 2016a), compared with HC. Activation results in the hippocampus tended to be mixed: for memory tasks HC subjects had increased activity for neutral or anxiety-induced conditions (Moon et al., 2015b; Moon et al., 2017) while subjects with GAD also had increased activity, but only for anxiety-induced conditions (Moon and Jeong, 2015b, 2017b; Moon et al., 2016). One conditioned fear task further showed increased activity for HC (Cha et al., 2016a), as well as for a generalized fear stimulus condition in a PPI FC study (Cha et al., 2014a). Finally, subjects with GAD showed increased FC with the hippocampus using dPFC (Wang et al., 2016a) and insula (Andreescu et al., 2015) seeds.

In addition to these four commonly accepted GAD-altered regions, a variety of other regions are also commonly altered. The insula, which has similar representation in the results as the hippocampus, appears to have reduced volume for subjects with GAD (Moon and Jeong, 2017a; Moon et al., 2014; Moon et al., 2015b), but greater FC (Buff et al., 2016; Fonzo et al., 2014; Liu et al., 2015; McClure et al., 2007; Qiao et al., 2017; Roy et al., 2013; Wang et al., 2016a)—particularly with amygdala seeds (Fonzo et al., 2014; Liu et al., 2015; McClure et al., 2007; Qiao et al., 2017; Roy et al., 2013). Only one result indicated reduced FC in the GAD insula (Andreescu et al., 2015). Insula activity was mixed, with greater activity in subjects with GAD for passive (Buff et al., 2016), congruency (Fonzo et al., 2014), and conditioned fear tasks (Laufer et al., 2016), mixed for emotion modulation tasks (reduced activity in Ball et al., 2013; and greater activity in Mohlman et al., 2017), and reduced in a prediction error task (White et al., 2017). The posterior cingulate cortex (PCC) is also fairly prevalent in the results, but has seldom been mentioned in previous reviews, and like the ACC tends to have mixed FC—greater in (McClure et al., 2007; Qiao et al., 2017; Strawn et al., 2012; Wang et al., 2016a) and reduced in (Etkin and Schatzberg, 2011; Qiao et al., 2017)—and task-based results, greater in (Buff et al., 2016; Fonzo et al., 2014; Mohlman et al., 2017) and reduced in (Carlson et al., 2017; Etkin and Schatzberg, 2011; Laufer et al., 2016; White et al., 2017), with no clear pattern emerging. Less common, but still each reported in at least 10 records, are the precuneus, precentral gyrus (largely from whole-brain analyses), superior
temporal gyrus, ventrolateral PFC (vlPFC), orbitofrontal cortex (OFC) and the cerebellum (supplementary Table S2).

The precuneus appears to have reduced FC with the dIPFC (Li et al., 2016; Wang et al., 2016a), mixed FC with the amygdala—greater in (McClure et al., 2007; Toazza et al., 2016) and reduced in (Strawn et al., 2012)—and reduced activity for working memory (Diwadkar et al., 2017; Moon and Jeong, 2015b, 2017b) in subjects with GAD. The precentral gyrus results show that FC tends to be greater, using amygdala (Monk et al., 2008; Toazza et al., 2016) and dIPFC (Wang et al., 2016a) seeds and activity is altered for working memory—greater in (Moon, 2015b) and reduced in (Moon et al., 2016, 2017)—reduced for a prediction error task (White et al., 2017), but increased for a condition fear task (Lauffer et al., 2016). Reduced volume is commonly, but not always observed in the precentral gyrus (Makovac et al., 2016a; Moon and Jeong, 2016, 2017a; greater volume in Strawn et al., 2013) and superior temporal gyrus (STG; greater volume in De Bellis et al., 2002; but reduced volume in Moon and Jeong, 2017a; Moon et al., 2014; Moon et al., 2015b) for GAD patients. Emotion modulation work resulted in decreased activity (Ball et al., 2013), while activity for condition fear (Lauffer et al., 2016) and FC (Liu et al., 2015; Monk et al., 2008; Roy et al., 2013; Wang et al., 2016a; Xia et al., 2017) was increased in the STG. The vIPFC showed reduced FA (Tromp et al., 2012) and increased FC (Andreescu et al., 2014; Li et al., 2018; Li et al., 2016; Monk et al., 2008; Roy et al., 2013), particularly using amygdala seeds (Li et al., 2016; Monk et al., 2008; Roy et al., 2013); however, decreased FC was also observed (Buff et al., 2016; Tromp et al., 2012 (amygdala seed)). Subjects with GAD had reduced activity for passive (Palm et al., 2011) and emotion modulation (Ball et al., 2013) tasks, greater activity for congruency (Monk et al., 2006) and memory tasks (Moon et al., 2015b; Moon et al., 2017), and mixed activity for conditioned fear tasks (reduced in Cha et al., 2016b; increased in Lauffer et al., 2016) in the vIPFC. The OFC has reduced mean diffusivity (Andreescu et al., 2017), cortical thickness (Andreescu et al., 2017), and surface area (Molent et al., 2017), mixed FC with prefrontal seeds, with greater FC in (Andreescu et al., 2015; Mohlman et al., 2017; Strawn et al., 2012; Wang et al., 2016a) and reduced FC in (Andreescu et al., 2015; Mohlman et al., 2017; Wang et al., 2016a). Additionally, the OFC has greater activity in subjects with GAD for emotion modulation (Mohlman et al., 2017; Paulsen et al., 2010) and passive (Fitzgerald et al., 2017) tasks, and reduced activity in conditioned fear (Lauffer et al., 2016) and memory (Diwadkar et al., 2017) tasks. Finally, whole-brain results show the midbrain is consistently smaller in subjects with GAD, as compared to HC (Moon and Jeong, 2015a, 2016, 2017a, 2017b; Moon et al., 2015b); however, these results are all from the same laboratory, and it is likely that there is some participant overlap between these records, although the authors could not be reached to confirm this.

The cerebellum results are again fairly mixed, having both increased (Andreescu et al., 2015; Liu et al., 2015; Roy et al., 2013) and reduced (Fonzo et al., 2014; Li et al., 2016; Roy et al., 2013) FC in subjects with GAD. However, grouping and re-labelling the results from the cerebellum yielded more distinct activation and FC patterns: HC > GAD contrasts were largely localized to the anterior lobe for FC (Li et al., 2016 [dIPFC seed]) and activity related to emotion regulation (Ball et al., 2013), congruency (Price et al., 2011), and working memory (Diwadkar et al., 2017; Moon and Jeong, 2015b, 2017b), with about half of the results localized to the culmen/vermis lobules IV and V (Ball et al., 2013; Li et al., 2016; Moon and Jeong, 2017b; see Table 3). Conversely, GAD > HC contrasts were largely observed in the posterior cerebellum with FC (Fonzo et al., 2014; Liu et al., 2015), and activity from congruency tasks (Fonzo et al., 2015; Monk et al., 2008; Price et al., 2011; see Table 3). Some papers in which cerebellum results were reported were excluded as specific contrasts were not done to compare subjects with GAD to HC (Benson et al., 2015; Brown et al., 2015; Carlisi et al., 2017; Haddad et al., 2015; Hamm et al., 2014; Lau et al., 2009; Park et al., 2016; Swartz et al., 2014). There is also at least one case in which cerebellum FC was hypothesized, but not observed (Toazza et al., 2016). As a caution to interpretation, the spatial accuracy of the cerebellum results may be limited as MNI or Talairach normalization can result in variability in fissure localization after registration—a SPM-compatible cerebellar atlas has been created for better spatial normalization in the future (Diedrichsen et al., 2009; Diedrichsen et al., 2011).

3.5. Meta-Analyses

Meta-analyses were conducted for VBM and task-based research (in which negative emotion-evoking tasks were compared to a neutral or null baseline). Records were excluded if they shared participants with another study—the record with the largest sample size was used. Whole-brain spmT maps were provided for two VBM records (Hilbert et al., 2015; Makovac et al., 2016a), and one task-based record (Price et al., 2011) while peak voxels were used in the remainder.

The VBM meta-analysis included six records (Hilbert et al., 2015; Liao et al., 2014b; Makovac et al., 2016a; Moon et al., 2014; Schienle et al., 2011; Strawn et al., 2013). Global volume changes could not be assessed between groups: only two records reported controlling for intracranial volume (Makovac et al., 2016a; Moon et al., 2014), but these values were only reported in one (Moon et al., 2014). GAD patients had greater volume than HC in several areas associated with visual processing (precuneus, angular, lingual, parahippocampal, fusiform, and middle occipital gyri), the inferior parietal gyrus, the pre- and postcentral gyri (Brodmann areas 1–4), the temporal pole and middle temporal gyrus. HC had greater volume than GAD along the cingulate cortex (cingulum, anterior cingulate/paracingulate), motor/planning regions (precentral gyrus [Brodmann area 6], supplementary motor area), and language areas (superior temporal gyrus [Heschl’s], inferior frontal gyrus, pars triangularis), and middle frontal gyrus (see Fig. 2 and supplementary Table S3).

The task-based meta-analysis was conducted with five records (Fitzgerald et al., 2017; Hözel et al., 2013; Monk et al., 2006; Palm et al., 2011; Price et al., 2011) in which authors conducted a between-subjects contrast, comparing visual, negative emotion-evoking stimuli with a neutral or null baseline. The tasks included making gender (Palm et al., 2011) or affect (Hözel et al., 2013) judgements of emotive faces, passively viewing or appraising images from the International Affective Picture System (IAPS; Fitzgerald et al., 2017), emotional Stroop task (Price et al., 2011), and distinguishing congruency with emotional faces (Monk et al., 2006). Peak or centre of gravity coordinates were not specified in one record (Monk et al., 2006), and contact with the authors revealed that this information could not be recalled. However, it was decided that because of the small size of the single, significant cluster in this record, that the coordinates would be included in the meta-analysis. GAD groups had greater activity in a cluster with the peak in the left amygdala (with additional local peaks, including in the striatum), the inferior network (uncinate fasciculus, orbital middle frontal gyrus), and the supramarginal gyrus, compared to HC groups for negative > neutral stimuli. Alternatively, the HC groups had greater activity in the orbital superior frontal gyrus (with additional local peaks throughout the middle frontal gyrus, and anterior cingulate/paracingulate gyri), and in the pars triangularis of the inferior frontal gyrus, compared to GAD groups for negative > neutral stimuli (see Fig. 2 and supplementary Table S4). Results from the jackknife sensitivity analyses can be observed in supplementary Table S5 for both meta-analyses. Although task-based results from the leave-one-out jackknife analyses tended to yield similar results—and when they differed, tended to result in clusters losing significance—one notable exception occurred when the record by Fitzgerald et al. (2017) was left out. For the GAD > HC contrast, a new, 104 voxel cluster in the cerebellum (hemispheric lobule 7, vermic lobules VI, VII, VIII, and crus I) was observed. These results should be regarded with caution as the Fitzgerald et al. (2017) record was, in fact, included, but may point to the need for further
Table 3
Cerebellum results across studies.

| Source | Normalization(WB or Seed) | Method | Contrast | Coordinates | Side | Lobe | Subregion |
|--------|---------------------------|--------|----------|-------------|------|------|-----------|
| **Healthy Control > Generalized Anxiety Disorder** |
| (Moon and Jeong, 2017b) | Talairach (WB) | Task—Recognition of faces after distractors | Neutral | 19 −32 −23 | R | A | Culmen (Vermis L. III) |
| (Ball et al., 2013) | Talairach (WB) | Task—Maintain or reduce reactions to images | Anxiety vs. Baseline | −37 −54 −24 | L | A | Culmen (Vermis L. III) |
| | | | Anxiety vs. Baseline | −34 −41 −20 | L | A | Culmen (Vermis L. III) |
| | | | Anxiety vs. Baseline | −26 −69 −28 | L | P | Pyramis (Vermis L. VII) |
| (Li et al., 2016) | MNI (R dlPFC) | Functional Connectivity—Resting state | Anxiety | −36 −56 −22 | L | P | Lobule 6 (Cerebellar H.) |
| (Moon and Jeong, 2015b) | MNI (WB) | Task—Recognition of faces after distractors | Neutral | 18 −34 −20 | R | A | Lobule 4, 5 (Cerebellar H.) |
| (Price et al., 2011) | MNI (WB) | Task—Emotional Stroop | Negative vs. Neutral | −22 −28 −24 | L | P | Lobule 6 (Cerebellar H.) |
| (Diwadkar et al., 2017) | MNI (WB) | Task—Memory/ suppression of word pairs | Suppression | 3 −43 −26 | R | A | N/A |
| | | | Retrieval | 3 −43 −26 | R | A | N/A |
| **Generalized Anxiety Disorder > Healthy Control** |
| (Fonzo et al., 2015) | Talairach (WB) | Task—Modified emotion face assessment task | Fear vs. happy | −2 −62 −36 | LR | P | Inf. Semi-Lunar Lobule (Crus II) |
| (Liu et al., 2015) | MNI (R Amygdala) | Functional Connectivity—Resting state | | −45 −63 −51 | L | P | Inf. Semi-Lunar Lobule (Crus II) |
| (Monk et al., 2009) | Talairach (WB) | Task—Congruency of neutral or emotional faces | Angry vs. neutral | 33 −30 −36 | R | P | Lobule 6 (Cerebellar H.) |
| (Andreescu et al., 2015) | MNI (L dlPFC) | Functional Connectivity—Worry perseverative cognition | | 6 −52 −2 | R | A | Clivus/Folium (Vermis L. IV, V) |
| (Fonzo et al., 2014) | Talairach (L Amygdala) | Functional Connectivity—PPI | | 8 −42 −21 | R | A | Clumun (Vermis L. III) |
| (Price et al., 2011) | MNI (WB) | Task—Emotional Stroop | Negative vs. Neutral | −2 −74 −22 | LR | P | Pyramis (Vermis L. VII) |

MNI regions were obtained by entering coordinates into MRicron software, and were labelled using the aal atlas overlay (https://www.nitrc.org/projects/mricron). Talairach regions were labelled by inputting coordinates into Talairach Client software (http://www.talairach.org/client.html). Although some records reported cerebellar activity within a cluster, if the peak results were outside of the cerebellum these results are not included here.

† These data were analysed in Montreal Neurological Institute (MNI) space, but results were converted to Talairach for reporting. WB = whole brain; MNI = Montreal Neurological Institute space; PPI = psychophysiological interaction; L = left; R = right; A = anterior; P = posterior; dlPFC = dorsolateral prefrontal cortex; Inf. = inferior; (Cerebellar H.) = cerebellar hemisphere; Vermis L. = Vermis Lobule.

Fig. 2. Results from the meta-analyses for GAD > HC (red) and GAD < HC (blue). Task-based results are for negative stimuli > neutral stimuli. See supplementary tables S3-4 for a full list of significant clusters. L = left; R = right; S = superior; I = inferior; A = anterior; P = posterior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
investigation.

4. Discussion

This systematic review and meta-analyses were concerned with determining the altered neural structure, FC, and activity in GAD patients. The current work makes an important contribution to the literature by providing corroborative evidence in support of the previously identified brain regions involved in GAD, and identifying novel brain regions not previously reported in systematic reviews. To our knowledge, this is the first meta-analytic investigation of GAD, as well as the largest systematic review to-date. This systematic review includes almost twice as many records as those included in any previous reviews and therefore provides the most current and comprehensive assessment of the neural correlates underlying GAD which furthers our understanding of this disorder.

The current systematic review, by using about twice as many additional studies and conducting two meta-analyses provides evidence for altered physiology in the dlPFC, ACC, amygdala, and hippocampus—three previous systematic reviews implicate these regions as well (Fonzo and Etkin, 2017; Hilbert et al., 2014; Mochcovitch et al., 2014). Interestingly, and importantly, these results, along with the others observed in the systematic review and meta-analyses lend themselves well to the idea of network-level organization—many of the altered regions are key structures in resting state networks. Although structure and function are largely related, structural metrics do not completely explain function (Batista-García-Ramo and Fernandez-Verdecia, 2018); for this reason this observation is speculative and exploratory, and it is important to note that structure, activity, and even FC alterations in these regions may not be directly related to resting state network FC or behavioural changes. Regardless, it remains interesting to consider the relationship between the implicated regions and their roles in network organization.

For example, the precuneus/PCC, medial PFC (mPFC), medial temporal lobes, and hippocampi are all nodes of the default mode network (Rosazza and Minati, 2011)—and all four of these regions had altered volume in the meta-analysis. Specifically we found increased volume in the middle temporal gyrus and precuneus and reduced volume in mPFC and hippocampus; reduced hippocampal volume was previously reported in one review (Hilbert et al., 2014). The default mode network is typically active during mind-wandering and self-referential thinking (Rosazza and Minati, 2011) and has often been observed as having altered FC in other psychopathologies (Broyd et al., 2009). Theoretical involvement of this resting state network in GAD makes sense as anxiety patients tend to ruminate with a self-referential focus (Broyd et al., 2009)—a key process attributed to this network. In another GAD systematic review, Fonzo and colleagues (Fonzo and Etkin, 2017) suggest that alterations of the anterior components of this network may be responsible for the “worry cascade” of GAD and that the worries formed in GAD are resistant to change because they seem to be immune to external, contradictory evidence.

The central executive (also known as the frontoparietal) network has almost the opposite role of the default mode network, being responsible for high-order cognitive processes such as maintaining objects in working memory, attention (Bressler and Menon, 2010), and coordinating cognitive control (Dixon et al., 2018; Marek and Dosenbach, 2018). This network appears pertinent to the GAD population from a behavioural perspective, likely manifested by difficulty concentrating, a common symptom in GAD. Further lending support to this idea are the brain nodes comprising this network: the dIPFC, inferior parietal gyrus (Sylvester et al., 2012), and crus II of the cerebellum (Shirer et al., 2012) have all been identified in our systematic review and the dIPFC and inferior parietal gyrus were also observed in the meta-analyses. In crus II, we observed increased FC between the right amygdala, and increased activity during the modified emotion face assessment task while our meta-analysis indicated greater volume in the inferior parietal cortex. Our results for the dIPFC were amongst the most prevalent: subjects with GAD had greater volume, and activity was mostly (but not entirely) reduced in response to passive, congruency, emotion modulation, and memory tasks. Additionally, FC tended to be reduced in the dIPFC, arising from amygdala, insula, and dIPFC seeds for GAD patients, although one study showed increased FC between the dIPFC and basolateral amygdala and another between the dIPFC and anterior insula. Previous GAD systematic reviews agree that PFC activity is altered (reduced in Mochcovitch et al., 2014) in subjects with GAD compared to HC (Fonzo and Etkin, 2017; Hilbert et al., 2014) for emotion regulation, and perseverative cognition. Hilbert and colleagues broke down the PFC results they observed by placing a larger emphasis on different age groups and found increased vlPFC activity for adolescents in attention/vigilance tasks, no differences in adults for an affective Stroop task, and increased dlPFC activity for neutral words, but decreased activity for negative words in an elderly GAD sample.

Because the default mode and central executive networks may have a role in GAD, it would be intuitive that the salience network may also be involved: this network is believed to act as a “switch” between the central executive and default mode networks (Shirer et al., 2012). The salience network is responsible for orienting attention to important (i.e., salient) information, and is thus implicated in threat-based responses—another indication that this network may be implicated in GAD. Interestingly, key nodes of the salience network—the ACC, insula, and amygdala (Bressler and Menon, 2010; Menon, 2015)—have been identified in the current systematic review and meta-analyses as regions likely being altered in GAD. Again the systematic review results for the ACC were mixed amongst a variety of tasks, corroborating previous reviews (Fonzo and Etkin, 2017; Hilbert et al., 2014; Mochcovitch et al., 2014). Fonzo and Etkin, (2017) address the variability in these results by concluding that the BOLD variability itself may be an intrinsic component of GAD, and that investigating the sources for this variability will be important for future understanding of this disorder. Although the ACC also had mixed FC results, overall they tended to be reduced for subjects with GAD when using an amygdala seed. Furthermore, meta-analyses showed reduced ACC activity and volume. The systematic review results for the amygdala indicated increased volume and FA for subjects with GAD, although our VBM meta-analysis failed to find volume differences in the amygdala—in line with a variety of studies failing to find expected amygdala results. Most of the task-based research indicated increased activity in GAD—including the task-based meta-analysis. Additionally, all three previous reviews (Fonzo and Etkin, 2017; Hilbert et al., 2014; Mochcovitch et al., 2014) discussed altered amygdala activity in GAD—sometimes hyperactivated for emotional stimuli only, sometimes hyperactivated for emotional and neutral stimuli, other times hypoactivated for fearful faces, and finally sometimes with no activity differences despite hypotheses to the contrary. Mochcovitch et al. (2014) suggested interpreting these amygdala results in tandem with the PFC response—especially because the reviews all highlight altered (reduced in Hilbert et al., 2014; Mochcovitch et al., 2014) FC between the amygdala and PFC. Because FC was reduced for GAD patients in dlPFC using amygdala and insula seeds, and in ACC using an amygdala seed—it seems likely that there may be some disconnection between the central executive and salience networks, which may contribute to or result from the idea that subjects with GAD have inflexibility in top-down processing (mediated by the default mode network), as mentioned by Fonzo and Etkin (2017).

Additionally, the sensorimotor network appears to have differences in many of its key nodes in GAD. The sensorimotor network includes the pre- and postcentral gyri, supplementary motor area (SMA), and cerebellum lobules IV/V/VI (Shirer et al., 2012): the meta-analyses indicates greater volume for subjects with GAD in the pre- and postcentral gyri, reduced volume in the SMA, and reduced activity in the cerebellum for tasks contrasting neutral and negative emotion-evoking stimuli. Although the systematic review shows mixed task-based results for the precentral gyrus for memory, fear learning, and prediction error.
tasks, the postcentral gyrus appears to have greater activity for subjects with GAD for fear learning, emotion modulation, and congruency tasks. As the sensorimotor network corresponds to the anatomy required for sensation and movement, and displays functionally relevant synchrony at rest (Rosazza and Minati, 2011), thus far, relation of this network to GAD remains speculative, but may be related to increased muscle tension and feelings of being “on edge” and hypervigilance in a motoric sense.

Delving deeper into the cerebellum, an often ignored region, there is a fairly substantial representation in the systematic review for FC and activity differences in GAD. Although initially, the results looked fairly mixed, running the cerebellum coordinates through Talairach Client or MRICron clarified the results. Compared to HC, GAD patients had reduced FC (largely with amygdala seeds) and activity in response to working memory, emotion modulation, and conflict tasks in the anterior lobe of the cerebellum (often in the culmen). Furthermore, compared to HC, GAD patients also had greater FC and activity for congruency and conflict, and facial affect processing tasks in the posterior cerebellum (Table 3). This anterior-posterior dichotomy becomes interesting in light of Bernard et al. (2012) assessment of the cerebellum FC. The authors found that the posterior cerebellar lobules correlated with prefrontal and association areas, indicating their involvement with the default mode network (Bernard et al., 2012)—it would be interesting to see if cerebellar and default mode networks had a stronger FC coupling since it appears that subjects with GAD have altered default mode and related cerebellar nodes.

Despite the relative lack of studies that report on the cerebellum, the idea of the cerebellum being altered in psychiatric disorders is not a new one: cerebellum volume or functional changes in psychiatric disorders including attention deficit hyperactivity disorder and schizophrenia has been observed (Baldacca et al., 2008; Phillips et al., 2015). Additionally, cerebellar volumes appear to be increased in OCD in the presence of childhood neglect (Brooks et al., 2016), while FC between the cerebellum and salience and executive control networks is altered in association with anxiety risk (Caulfield et al., 2016).

A recent consensus paper by Adamaszek et al. (2017) indicates that in addition to its well-known role in regulating motor control, the cerebellum also plays a role in a wide variety of emotion processing. The culmen specifically (vermis lobules IV/V) has been shown to be hypoactive in alexithymia—a condition marked by dysfunctional emotional awareness (Adamaszek et al., 2017). Adamaszek also reported on a meta-analysis implicating vermal lobules IV and VI in explicit emotional face processing (Adamaszek et al., 2017). The inferior semi-lunar lobules (cerebellar hemisphere VIIB) have been shown to be active in response to unpleasant images when combined with noxious heat (Adamaszek et al., 2017). Although a clear picture is emerging for the localization of cerebellar alteration in GAD, the roles that each region plays remains complex as they appear to be involved in emotion-related processing, in addition to the better-known roles of motor control.

This review and meta-analysis all tend to point towards the same conclusion of the previous reviews: top-down, emotion dysregulation and the localization of cerebellar alteration in GAD, the roles that each region plays remains complex as they appear to be involved in emotion-related processing, in addition to the better-known roles of motor control.

5. Limitations

A number of limitations exist within the present work. First, this review is limited in that only studies employing direct comparisons between GAD and HC were included. Furthermore, differences between GAD patients and additional disorders were largely ignored to maintain the focus of the systematic review. Finally, the meta-analyses performed were limited in terms of the number of records eligible for inclusion, and the availability of whole-brain maps. Although many authors were more than willing to share their data, in many cases, data loss resulting from technical limitations and maintaining ethics requirements, in addition to other hindrances, greatly limited access to whole-brain data. The resulting sample size for each of the meta-analyses further limited the complementary analyses that could be conducted, resulting in a mixture of population ages, medication use, and comorbidities. Finally, although many of the regions identified in the systematic review and meta-analyses are key nodes of resting state networks, it is important to note that many of these results are structural or activity-based in nature and may not as clearly relate to or affect the function of whole-brain resting state networks themselves—future whole-brain resting state studies of GAD can help to further investigate this.

6. Conclusion

This review summarizes a large body of work focusing on the neural underpinnings of GAD and has produced strong evidence for the involvement of specific brain regions. Previously accepted altered regions include the dIPFC ([‘] indicate meta-analysis results while no brackets indicate systematic review results: [reduced volume], altered FC with amygdala, altered [reduced] activity), ACC ([reduced volume], mixed FC and mixed [reduced] activity), amygdala (increased [increased] volume, increased activity), and hippocampus (greater left-lateralized volume) in the GAD literature. Additionally, previously unidentified regions including the insula (reduced volume, greater FC, mixed [greater] activity for GAD), PCC (increased volume), mixed FC, and mixed [increased] activity), precuneus ([increased volume], altered FC, reduced working memory activity), precentral gyrus (reduced [increased in right, increased in left hemisphere] volume, greater FC, mixed activity), STG (reduced [in reduced in left, greater right] volume, increased FC, [increased activity]), vlPFC ([reduced volume], mostly increased FC, mixed [reduced] activity), OFC (mean diffusivity, cortical thickness and surface area, mixed FC and mixed [reduced] activity), and cerebellum (reduced FC and working memory activity in anterior lobe, greater FC and congruency-based activity in posterior cerebellum, reduced activity) are identified as regions of interest via both our systematic review and our meta-analyses. Despite the use of different modalities (i.e., structure, FC, and task-based methods) and widely varying methods of analyses within each modality (e.g., VBM vs. FA values)—a high degree of consistency was observed within the systematic review and meta-analyses. This consistency was observed despite a high degree of variability in terms of age groups, comorbidities, and medication use included in each record. Future research should be conducted to determine if and how these regions differ with severity and duration of the disorder, and between different mood and anxiety disorders. Through this process, we may begin to better understand how the alterations in neural structures and networks contribute to the development and/or maintenance of GAD, which may in turn inform treatment strategies for this patient population.

CRediT authorship contribution statement

Tiffany A. Kolesar: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Writing - original draft, Writing - review & editing. Elena Bilevicius: Conceptualization, Data curation, Writing - review & editing. Alyssia D. Wilson: Data curation, Writing - review & editing. Jennifer Kornelsen: Resources, Software, Supervision, Writing - review & editing.

Funding

This work was supported by the Natural Sciences and Engineering
Research Council of Canada (NSERC), Canada Graduate Scholarships Doctoral Program.

Declaration of Competing Interest
None.

Supplementary material
Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2019.102016.

References

Abdollahi, C.G., Coplan, J.D., Jackowski, A., Sato, J.R., Mao, X., Shungu, D.C., Mathew, S.J., 2013. A pilot study of hippocampal volume and n-acetylaspartate (NAA) as response biomarkers in ritilin-treated patients with GAD. Eur. Neuropsychopharmacol. 23 (4), 276–284. https://doi.org/10.1016/j.euroneuro.2012.05.009.

Adamaszek, M., D’Agata, F., Ferrucci, R., Habas, C., Kirby, K.C., Verhoeven, J., 2017. Consensus paper: cerebellum and emotion. Cerebellum 16 (2), 552–576. https://doi.org/10.1007/s12311-016-0815-8.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. American Psychiatric Association Publishing, Arlington, VA.

Andreescu, C., Gross, J.J., Enote, E., Edelman, K.D., Snyder, S., Tanase, C., Alizentin, H., 2011. Altered cerebral blood flow patterns associated with pathologic worry in the elderly. Depress Anxiety 28 (3), 202–209. https://doi.org/10.1002/da.20799.

Andreescu, C., Sheu, L.K., Tudorascu, D., Gross, J.J., Walker, S., Banhaishemi, L., Alizentin, H., 2015. Emotion reactivity and regulation in late-life generalized anxiety disorder: functional connectivity at baseline and post-treatment. Am. J. Geriatr. Psychiatry 23 (2), 200–214. https://doi.org/10.1016/j.jagp.2014.05.003.

Andreescu, C., Sheu, L.K., Tudorascu, D., Walker, S., Alizentin, H., 2014. The age of anxiety - Differences across the lifespan in the default mode network functional connectivity in generalized anxiety disorder. Int. J. Geriatr. Psychiatry 29 (7), 704–712. https://doi.org/10.1002/gps.4051.

Andreescu, C., Tudorascu, D., Sheu, L.K., Raagharajam, A., Butters, M.A., Walker, S., Alizentin, H., 2017. Brain structural changes in late-life generalized anxiety disorder. Psychiatry Res.-Neuroimaging 268, 15–21.

Baldacara, L., Borgio, J.G., Lacerda, A.L., Jackowski, A.P., 2008. Cerebellum and psychiatric disorders. Rev. Bras. Psiquiatr. 30 (3), 281–289.

Ball, T.M., Ramsawh, H.J., Campbell-Sills, L., Paulus, M.P., Stein, M.B., von Kienlin, M., Liu, T.T., Simmons, A., Paulus, B.E., Guyer, A.E., Nelson, E.E., Pine, D.S., Ernst, M., 2015. Role of contingency in anxiety disorder: functional connectivity at baseline and post-treatment. Am. J. Geriatr. Psychiatry 23 (3), 276–284. https://doi.org/10.1016/j.jagp.2014.05.003.

Cha, J., Carlson, L.M., DeDora, D.J., Greenberg, T., Proudfit, G.H., Mujica-Parodi, L.R., 2014a. Hyper-reactive human ventral temporal area and aberrant mesocorticolimbic connectivity in overgeneralization of fear in generalized anxiety disorder. J. Neurosci. 34 (17), 5885–5890. https://doi.org/10.1523/JNEUROSCI.4868-13.2014.

Cha, J., DeDora, D., Nedic, S., Ide, J., Greenberg, T., Hajcak, G., Mujica-Parodi, L.R., 2014b. Clinically anxious individuals show disrupted feedback between the lateral prefrontal cortex and ventromedial prefrontal-limbic control circuit. J. Neurosci. 34 (17), 4708–4718. https://doi.org/10.1523/JNEUROSCI.1092-15.2016.

Cha, J., Greenberg, T., Carlson, L.M., DeDora, D.J., Hajcak, G., Mujica-Parodi, L.R., 2015. Circuit-wide structural and functional measures predict ventromedial prefrontal cortex fear generalization: implications for generalized anxiety disorder. J. Neurosci. 34 (41), 4043–4053. https://doi.org/10.1523/JNEUROSCI.3372-13.2014.

Cha, J., Greenberg, T., Song, I., Blair Simpson, H., Pomer, J., Mujica-Parodi, L.R., 2016b. Abnormal hippocampal structure and function in clinical anxiety and comorbid depression. Hippocampus 26 (5), 545–553. https://doi.org/10.1002/hipo.22566.

Chen, A.C., Etkin, A., 2013. Hippocampal network connectivity and activation differentiates post-traumatic stress disorder from generalized anxiety disorder. Neuropsychopharmacology 38 (10), 1889–1898. https://doi.org/10.1038/npp.2013.122.

Cui, H., Zhang, J., Liu, Y., Li, Q., Li, H., Zhang, I., Northoff, G., 2016. Differential alterations of resting-state functional connectivity in generalized anxiety disorder and panic disorder. Hum. Brain Mapp. 37 (4), 1459–1472. https://doi.org/10.1002/hbm.23113.

De Bellis, M.D., Casey, B.J., Dahl, R.E., Birnbaum, B., Williamson, D.E., Thomas, K.M., Ryan, N.D., 2000. A pilot study of amygdala volumes in pediatric generalized anxiety disorder. Biol. Psychiatry 48 (1), 51–57. https://doi.org/10.1016/s0006-3223(00)00835-0.

De Bellis, M.D., Keshavan, M.S., Shifflett, H., Iyengar, S., Dahl, R.E., Axelson, D.A., Ryan, N.D., 2002. Superior temporal gyri volumes in pediatric generalized anxiety disorder. Biol. Psychiatry 51 (7), 553–562. https://doi.org/10.1016/s0006-3223(01)01375-0.

Diedrichsen, J., Balsters, J.H., Havell, J., Cussans, E., Ramnani, N., 2009. A probabilistic MR atlas of the human cerebellum. Neuroimage 46 (1), 39–49. https://doi.org/10.1016/j.neuroimage.2009.01.045.

Diedrichsen, J., Maderwald, S., Kuper, M., Thurling, M., Rabe, K., Gizewski, E.R., Timmmann, I. 2011. Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. Neuroimage 54 (3), 1786–1794. https://doi.org/10.1016/j.neuroimage.2010.10.035.

Diwadkar, V.A., Re, M., Cecchetto, F., Garzotto, M., Piccin, S., Bonventito, C., Brambilla, P., 2017. Attempts at memory control induce dysfunctional brain activation profiles in generalized anxiety disorder: an exploratory fMRI study. Psychiatry Res.-Neuroimaginq 266, 42–52. https://doi.org/10.1016/j.pscychresns.2017.05.010.

Dixon, M.L., De La Vega, A., Mills, C., Andrews-Hanna, J., Spreng, R., Cole, M.W., Christoff, K., 2018. Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. Proc. Natl. Acad. Sci. U S A 115 (7), E1598–E1607. https://doi.org/10.1073/pnas.1715766115.

Duval, E.R., Javanbakht, A., Liberson, I., 2015. Neural circuits in anxiety and stress disorders: a focused review. Ther. Clin. Risk Manag. 11, 115–126. https://doi.org/10.2147/tcrm.s74852.

Etkin, A., Prater, K.E., Hoenf, F., Menon, V., Schatzberg, A.F., 2010. Failure of anterior cingulate activation and connectivity with the amygdala during implicit regulation of emotional processing in generalized anxiety disorder. Am. J. Psychiatry 167 (5), 551–554. https://doi.org/10.1176/appi.ajp.2009.09104.

Etkin, A., Prater, K.E., Schatzberg, A.F., Menon, V., Greicius, M.D., 2009. Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder. Arch. Gen. Psychiatry 66 (12), 1361–1372. https://doi.org/10.1001/archgenpsychiatry.2009.104.

Etkin, A., Schatzberg, A.F., 2011. Common abnormalities and disorder-specific compensation during implicit regulation of emotional processing in generalized anxiety and major depressive disorder. Am. J. Psychiatry 168 (9), 968–978. https://doi.org/10.1176/appi.ajp.2011.1009290.

Fitzgerald, J.M., Phan, K.L., Kennedy, A.E., Shankman, S.A., Langenecker, S.A., Klump, H., 2017. Prefrontal and amygdala engagement during emotional reactivity and regulation in generalized anxiety disorder. J. Affect. Disord. 218, 398–406. https://doi.org/10.1016/j.appseer.2016.12.002.
