Journal of Psychiatric Research 68 (2015) 91–98

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Subcortical volumes differentiate Major Depressive Disorder, Bipolar Disorder, and remitted Major Depressive Disorder

Matthew D. Sacchet [a, b, *], Emily E. Livermore [b], Juan Eugenio Iglesias [c], Gary H. Glover [a, d, e], Ian H. Gotlib [a, b]

[a] Neurosciences Program, Stanford University, Stanford, CA, USA
[b] Department of Psychology, Stanford University, Stanford, CA, USA
[c] Basque Center on Cognition, Brain and Language, San Sebastian – Donostia, Gipuzkoa, Spain
[d] Department of Radiology, Stanford University, Stanford, CA, USA
[e] Department of Biophysics, Stanford University, Stanford, CA, USA

ARTICLE INFO

Article history:
Received 29 April 2015
Accepted 5 June 2015

Keywords:
Bipolar Disorder (BD)
Major Depressive Disorder (MDD)
Remitted Major Depressive Disorder (RMD)
Ventral diencephalon
Caudate nucleus
Subcortical

ABSTRACT

Subcortical gray matter regions have been implicated in mood disorders, including Major Depressive Disorder (MDD) and Bipolar Disorder (BD). It is unclear, however, whether or how these regions differ among mood disorders and whether such abnormalities are state- or trait-like. In this study, we examined differences in subcortical gray matter volumes among euthymic BD, MDD, remitted MDD (RMD), and healthy (CTL) individuals. Using automated gray matter segmentation of T1-weighted MRI images, we estimated volumes of 16 major subcortical gray matter structures in 40 BD, 57 MDD, 35 RMD, and 61 CTL individuals. We used multivariate analysis of variance to examine group differences in these structures, and support vector machines (SVMs) to assess individual-by-individual classification. Analyses yielded significant group differences for caudate ($p = 0.029$) and ventral diencephalon (VD) volumes ($p = 0.003$). For the caudate, both the BD ($p = 0.004$) and the MDD ($p = 0.037$) participants had smaller volumes than did the CTL participants. For the VD, the MDD participants had larger volumes than did the BD and CTL participants ($ps < 0.005$). SVM distinguished MDD from BD with 59.5% accuracy. These findings indicate that mood disorders are characterized by anomalies in subcortical gray matter volumes and that the caudate and VD contribute uniquely to differential affective pathology. Identifying abnormalities in subcortical gray matter may prove useful for the prevention, diagnosis, and treatment of mood disorders.

1. Introduction

Mood disorders are among the most prevalent and severe of all psychiatric disorders (Kessler et al., 2005; WHO, 2012). Whereas both Major Depressive Disorder (MDD) and Bipolar Disorder (BD) are characterized by the presence of depressive episodes, BD is also associated with manic or hypomanic episodes. Because BD often presents clinically as a depressive episode, patients experiencing this disorder can be misdiagnosed as MDD, leading to inappropriate treatment and prolonged distress (Singh and Rajput, 2006). We know little about neurobiological differences between BD and MDD (de Almeida and Phillips, 2013), which limits effective prevention, diagnosis, and treatment of these disorders.

Subcortical gray matter structures are involved in cognitive processing and emotion generation and regulation (Lindquist et al., 2012; Ochsner et al., 2012); not surprisingly, therefore, investigators have implicated anomalies in these structures in mood disorders (Savitz and Drevets, 2009). More specifically, individuals diagnosed with mood disorders have been found to be characterized by structural and functional abnormalities in the amygdala, hippocampus, caudate and putamen, pallidum, nucleus accumbens, and thalamus (Savitz and Drevets, 2009; Hamilton et al., 2012).

Using meta-analytic methods, Kempton et al. compared regional brain volumes in MDD and BD participants and found that the caudate, corpus callosum cross-sectional area, putamen, pallidum,

* Corresponding author. Department of Psychology, Jordan Hall, Building 01-420, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA. Tel.: +1 650 723 0804.
E-mail addresses: msacchet@stanford.edu (M.D. Sacchet), eliv@stanford.edu (E.E. Livermore), e.iglesias@bcbl.eu (J.E. Iglesias), gary.glover@stanford.edu (G.H. Glover), ian.gotlib@stanford.edu (I.H. Gotlib).

http://dx.doi.org/10.1016/j.jpsychires.2015.06.002
0022-3956/

and hippocampus are smaller in MDD than in BD (Kempton et al., 2011). Importantly, these results were limited to common brain regions previously studied in both MDD and BD, and are susceptible to biases resulting from a wide range of participant inclusion criteria and neuroimaging and statistical methods across studies. Moreover, the broad comparisons of MDD versus BD did not account for heterogeneous disease states, including influences from BD I and BD II, euthymic, manic, hypomanic and depressed BD, and current and remitted MDD. Finally, Kempton et al.'s results may be confounded by differences in illness severity between MDD and BD. It is important, therefore, that investigators directly compare MDD and BD individuals in different states with comparable illness history.

To date, few studies have examined differences in brain structure between individuals diagnosed with MDD and BD. The results of these studies indicate that, compared to MDD, BD is associated with greater deep white matter hyperintensities (Dupont et al., 1995; Silverstone et al., 2003), reduced fractional anisotropy of the left superior longitudinal fasciculus (Versace et al., 2010), decreased habenula volume (Savitz et al., 2011), reduced cortical thickness in caudal middle frontal cortex, inferior parietal cortex, and precuneus (Lan et al., 2014), and increased thalamic volume (Dupont et al., 1995). In addition, Redlich et al. found clusters of reduced gray matter that spanned the hippocampal formation, amygdala, putamen, insula, and temporal pole in depressed BD compared to MDD individuals, and a cluster in anterior cingulate that was reduced in MDD compared to depressed BD individuals (Redlich et al., 2014).

Recently, investigators have begun to examine characteristics of MDD and BD that may persist beyond the clinical episode of depression or mania. For example, researchers have found that individuals with BD who are currently in remission exhibit impairment on tests of visuospatial recognition memory (Rubinsztein et al., 2000). Similarly, in a review of studies of cognitive impairment in individuals who had recovered from MDD, Hasselbach et al. (2011) found that in 9 of 11 of these studies remitted depressed participants exhibited impaired performance on at least one neuropsychological test (Hasselbalch et al., 2011). Researchers have also found that individuals continue to experience impairment in social and occupational functioning following remission of MDD or BD (e.g., Fagiolini et al., 2005; Romera et al., 2010). Importantly, investigators have documented abnormalities in regional brain volumes in individuals who have remitted from MDD and BD. For example, individuals with euthymic BD have lower metabolic rates than do healthy controls and depressed BD individuals (Yildiz et al., 2001). Similarly, individuals with remitted MDD have smaller total and posterior hippocampal volumes than do healthy controls (Neumeister et al., 2006; for review see Lorenzetti et al., 2009). Understanding temporary (i.e., state) vs. enduring (i.e., trait) characteristics of affective disorders will facilitate the identification of targets for prevention and treatment. This is particularly important for MDD and BD, given that improved characterization of remitted MDD and euthymic BD may allow for greater differentiation of these topographically similar states and help to avoid maladaptive consequences of misdiagnosis (Singh and Rajput, 2006).

In this study we directly compare, for the first time, subcortical volumetric differences between individuals diagnosed with BD who are currently euthymic and individuals diagnosed with MDD. In addition, to examine the state versus trait nature of volumetric anomalies in mood disorders, we included a sample of individuals with remitted Major Depression (RMD), in addition to a group of healthy (CTL) individuals. We used FreeSurfer's automated segmentation to assess regional subcortical gray matter volumes of the accumbens area, amygdala, caudate, hippocampus, pallidum, putamen, thalamus, and ventral diencephalon (VD; including hypothalamus). To assess the relation of volumetric abnormalities to the severity of impairment in data-to-day functioning across disorders, we related these volumes to individuals' level of global functioning (Global Assessment of Functioning [GAF]; Endicott et al., 1976). Finally, we used support vector machines (SVMs) to examine whether identified abnormal volumes can be used to classify participants on an individual-by-individual basis (Cortes and Vapnik, 1995).

We hypothesized that MDD individuals would have smaller volumes than would BD and CTL individuals in the regions identified in Kempton et al.'s meta-analysis, including caudate, pallidum, putamen and hippocampus. In addition, based on Redlich et al.'s findings with currently depressed BD individuals (Redlich et al., 2014), we hypothesized MDD-related reductions in amygdala relative to BD individuals. Although Kempton et al. did not find significant differences between BD and CTL participants in these regions, Redlich et al. found BD-related abnormalities that spanned hippocampus, amygdala, caudate, putamen, and thalamus; thus, we hypothesized that BD individuals would be distinguishable from CTLs in these regions. We also hypothesized that volumes of the RMD participants would fall between those of MDD and BD, and MDD and CTL participants. Finally, we hypothesized that using SVMs, the identified abnormal regions would successfully classify the MDD versus BD and both the MDD and BD versus CTL groups.

2. Materials and methods

2.1. Participants and clinical information

Participants were 193 individuals: 40 diagnosed with BD (Bipolar I Disorder, all currently euthymic); 57 diagnosed with MDD; 35 diagnosed with past but not current MDD (RMD); and 61 CTLs. All individuals participated in studies at Stanford University in which MRI data were acquired. The Structured Clinical Interview for DSM was administered by trained interviewers to all participants in order to obtain DSM-IV-TR Axis I diagnoses (First et al., 2004). Our team of interviewers have demonstrated high inter-rater reliability in our samples for these diagnoses ($\kappa s > 0.9$; e.g., Levens and Gotlib, 2010; Victor et al., 2011; Johnson et al., 2012). No participant met diagnostic criteria for substance or alcohol abuse or dependence within six months prior to MRI scanning. CTL individuals did not meet diagnostic criteria for any current psychiatric disorder or past mood disorder. Interviewers also assessed level of global functioning, using the GAF scale (Endicott et al., 1976) and number of lifetime Major Depressive episodes (MDEs) and lifetime manic episodes. Scores on the GAF scale range from 1 to 100 (lowest to highest level of functioning), indexing individuals' level of occupational, psychological, and social functioning. Participants in all four groups were assessed with the GAF; relating this measure to volumetric abnormalities might offer insight into how such abnormalities are related to the day-to-day functioning of individuals with affective disorders. Written informed consent was obtained from each participant; the Stanford University Institutional Review Board approved the study.

2.2. MRI data acquisition

All data were collected using the same 1.5 T magnetic resonance imaging (MRI) system and no major scanner upgrades that would influence SPGR images were undertaken. Further details are included in the Supplemental Information.

2.3. Subcortical segmentation

We used the FreeSurfer software suite for the automated segmentation of subcortical volumes from the T1w images (Fischl et al., 2002). The FreeSurfer automatic segmentation has been shown to be reliable and to have accuracy comparable to that of manual labeling techniques (Fischl and Dale, 2000; Fischl et al., 2002), please see the Supplemental Information for additional details concerning FreeSurfer reliability. The FreeSurfer automated segmentation has been widely applied to measure subcortical regional volumes. We analyzed data from left and right accumbens area, amygdala, caudate, hippocampus, pallidum, putamen, thalamus, and VD. For each region, we regressed out the following covariates of non-interest: age at scan, gender, and estimated intracranial volume (ICV). All subsequent analyses were conducted on the resulting model residuals. Further details are included in the Supplemental Information.

2.4. Assessing group differences in volume

A two-way (group repeated over hemisphere) multivariate analysis of variance (MANOVA) was conducted using the SPSS Statistics software package (IBM Corporation, Armonk, USA) to identify regional subcortical volumetric differences among the four diagnostic groups (BD, MDD, RMD, and CTL). Significant multivariate effects were followed up with appropriate univariate analyses. Finally, after identifying neural regions for which there were significant group differences, we subjected volumes of these regions to individual-by-individual classification using machine learning, specifically, linear SVMs (Cortes and Vapnik, 1995) with and without Recursive Feature Elimination (RFE; De Martino et al., 2008). Details concerning our machine learning methods are presented in the Supplemental Information.

2.5. Relation of subcortical volumes to level of global functioning

For each region that significantly differentiated the four groups in the MANOVA, we computed Pearson correlation coefficients to assess the relation between volume in the region and scores on the GAF.

3. Results

3.1. Demographic and clinical characteristics

Demographic and clinical characteristics of the participants in the four groups are presented in Table 1, medications are presented in Supplement Table S2, and comorbidity information is presented in Supplement Table S3. The RMD group had significantly more female ($\chi^2(3) = 12.186$, $p = 0.007$) and older ($F(3,189) = 3.076$, $p = 0.029$) participants than did the other three groups, who did not differ significantly from each other (LSD post-hoc tests and two-group chi-square tests $p > 0.05$). One CTL and one BD individual did not have GAF scores, and two MDD and four BD individuals did not have number of lifetime MDE information. The four groups also differed significantly in GAF scores ($F(3,187) = 152.893$, $p < 0.001$); as expected, participants in the CTL group had the highest GAF scores, followed by the RMD, BD, and MDD groups (all LSD post-hoc tests $p < 0.05$). The four groups did not differ with respect to ICV ($F(3,189) = 0.9$, $p = 0.442$), and the three clinical groups did not differ in the number of lifetime MDEs ($F(2,122) = 1.7$, $p = 0.195$).

3.2. Subcortical volumetric differences among groups

A two-way (group repeated over hemisphere) MANOVA conducted on the eight regional volume residuals in each hemisphere yielded a significant main effect for group ($F(24, 528.5) = 1.634$; $p = 0.03$; Wilks' $\Lambda = 0.812$; partial $\eta^2 = 0.067$); neither the main effect for hemisphere nor the interaction of group and hemisphere was significant ($ps > 0.1$). Follow-up univariate (by group) repeated-measures (over hemisphere) analyses of variance (ANOVAs) conducted for each region yielded significant effects for group for volumes of the caudate ($F(3, 189) = 3.081$, $p = 0.029$; partial $\eta^2 = 0.047$) and VD ($F(3, 189) = 4.832$, $p = 0.003$; partial $\eta^2 = 0.071$; see Fig. 1 panels C and D). LSD post-hoc tests indicated that for caudate volumes, the CTL group had significantly larger volumes than did the BD ($p = 0.004$) and MDD ($p = 0.037$) groups (Fig. 2 panel A). For the VD, the MDD group had significantly larger volumes than did the BD ($p = 0.001$) and CTL ($p = 0.004$) groups, and the RMD group had marginally larger volumes than did the BD group ($p = 0.068$; Fig. 2 panel B).

We used SVMs to assess whether regions that were identified in the MANOVA could be used to classify group membership. Using left and right caudate and VD volume residuals, we were able to classify the MDD vs. CTL (average accuracy across subsamples = 62.44% without RFE and 62.76% with RFE; $ps < 0.01$) and the MDD vs. BD (average accuracy across subsamples = 59.45% with RFE; $p < 0.05$) groups on an individual-by-individual basis. See the Supplemental Information (Table S4) for additional findings, including complete classification results and measures related to sensitivity and specificity.

3.3. Relation between subcortical volumes and GAF scores

To assess the relation between subcortical volumes that were identified in the ANOVAs and participants' global level of

Table 1
Demographic and clinical information. The remitted depressed (RMD) group was older and had more females than did the other groups. The four groups differed in Global Assessment of Functioning (GAF) scores (one CTL and one BD individual did not have GAF scores): participants in the CTL group had the highest GAF scores, followed by the RMD, Bipolar Disorder (BD), and Major Depressive Disorder (MDD) groups.

	BD ($N = 40$)		MDD ($N = 57$)		RMD ($N = 35$)		CTL ($N = 61$)		p-value
Age in years (yrs; $M\|SD$)	37.8_a	9.6	37.1_a	10.1	42.9	8.6	37.2_a	10.4	0.029
Gender (count F|% F)	21_a	52.5	34_a	59.7	31	88.6	37_a	60.7	0.007
ICV (1^*10^{-6} mm^3 $M\|SD$)	1.46	0.13	1.42	0.16	1.41	0.14	1.45	0.17	0.442
GAF ($M\|SD$)	70.0	13.4	51.7	7.8	75.4	10.4	87.3	5.0	<0.001
Total number of lifetime MDEs*	10.7	12.8	12.1	10.6	8.1	8.6			0.195
Total number of lifetime manic episodes	8.8	11.1							NA
Duration of current MDE (months)			12.1	15.0					NA

M = mean; SD = standard deviation; MDE = Major Depressive episode; BD = Bipolar Disorder group; MDD = Major Depressive Disorder group; RMD = Remitted Major Depressive group; CTL = healthy control group; yrs = years; F = female; ICV = intracranial volume; GAF = Global Assessment of Functioning scale; * extreme values (e.g., participant responded "too many to count") were set to 25; Age in years, ICV, GAF, and total number of lifetime MDEs were compared across groups using one-way ANOVA; Gender was compared across groups using chi-square; groups with the same subscript within rows do not differ from each other at $p < 0.05$ as computed using LSD post-hoc tests for ANOVA and two-group chi-square tests.

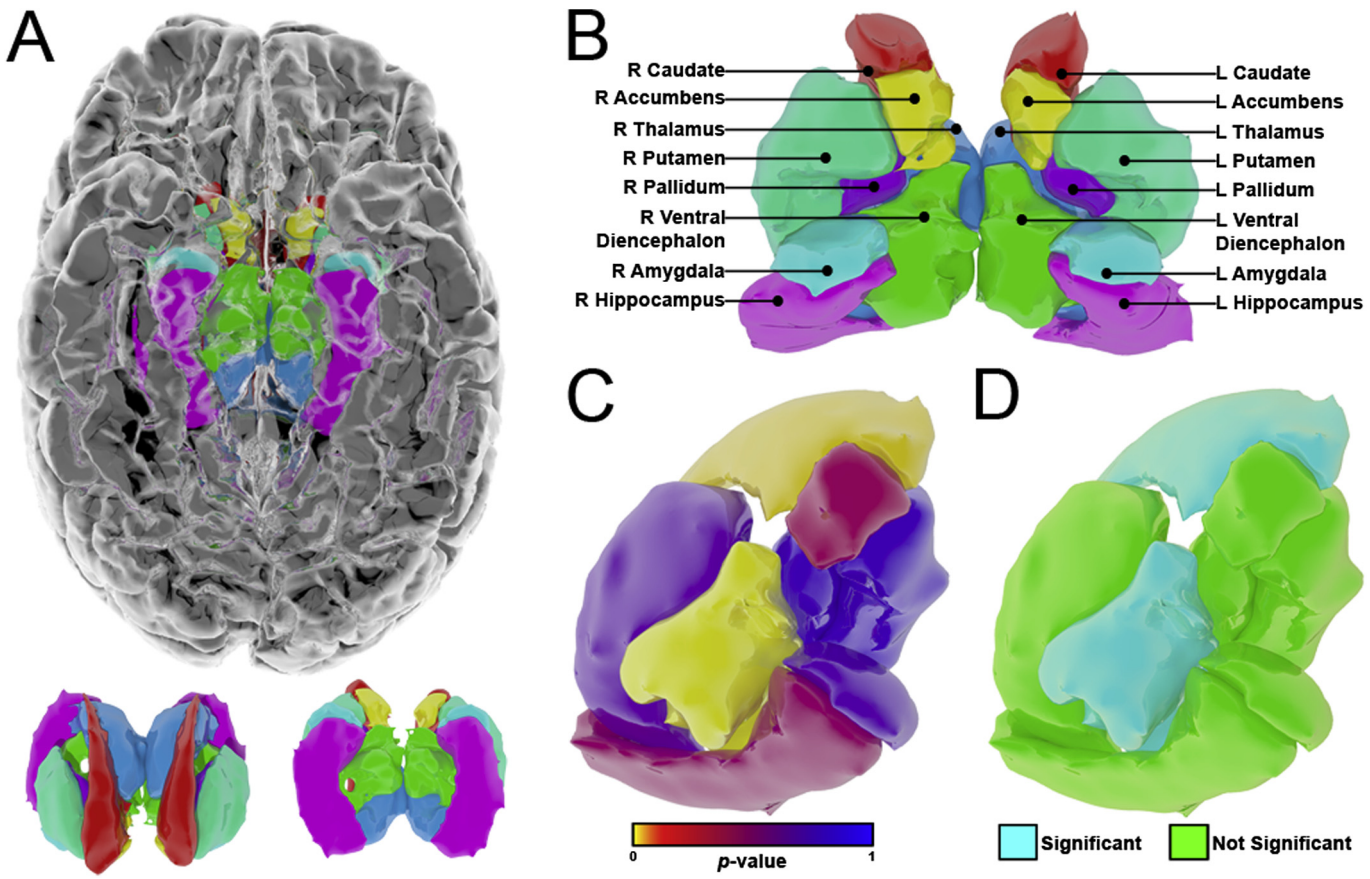

Fig. 1. Subcortical regions of interest (ROIs) and regional analysis of variance (ANOVA) results. A) whole brain image from ventral orientation with transparent brain and colored subcortical structures and panels with superior and ventral views of subcortical structures. B) Subcortical structures with labels. C) *p*-values for one-way analysis of variance (ANOVA) for each region over groups. D) Subcortical regions with statistical significance of ANOVAs indicated: caudate ($p = 0.029$) and ventral diencephalon ($p = 0.003$).

functioning, we computed correlations between GAF scores and mean caudate volume residuals and mean VD volume residuals within each of the three disordered groups. GAF scores were found to be positively correlated with VD volume within the MDD group ($r = -0.294$; $p = 0.026$) and negatively correlated with caudate volume within the BD group ($r = -0.385$; $p = 0.014$).

4. Discussion

The goal of this study was to identify subcortical volumes that differentiate participants diagnosed with MDD and BD, RMD participants, and healthy CTL participants, and to use the identified subcortical brain volumes to classify disorder status on an individual-by-individual basis. We found that the BD and MDD groups had smaller caudate volumes than did the CTL group, and that the VD was larger in the MDD group than in the BD and CTL groups, and marginally larger in the RMD than in the BD group. GAF scores were inversely correlated with caudate volume in the BD group and with VD volume in the MDD group, suggesting that these volumetric abnormalities are related to the severity of impairment in day-to-day functioning. Finally, using SVMs, we were able to differentiate the MDD from the CTL and BD groups.

Although traditionally the basal ganglia (composed of the caudate, putamen, pallidum, subthalamic nucleus, and substantia nigra) were conceptualized as a central component of motor processes, increasing evidence suggests that these regions play a role in affective processing (Savitz and Drevets, 2009). Indeed, the caudate is one of the two components of the striatum, which is a critical component of the cortico-striato-pallido-thalamic (CSPT) and amygdalo-striato-pallido-thalamic (ASPT) loops. These circuits have been posited to form the core of the neural systems associated with abnormality in mood disorders (Price and Drevets, 2009, 2012). Specifically, these neural systems are involved in reward-seeking behaviors including reward anticipation and evaluation (Kawagoe et al., 1998; Gold, 2003). Importantly, a core symptom of depression is anhedonia, or the inability to experience pleasure; consequently, the CSPT and ASPT have been of considerable interest to researchers studying this disorder. Our observation in the present study of reduced caudate volumes in MDD and BD suggests that this structure is implicated in reward-related abnormalities in affective disorders.

Studies have been inconsistent in their findings of caudate volumetric abnormalities in MDD. Whereas several researchers have reported reduced caudate volumes in individuals with MDD (e.g., Kim et al., 2008 [although there was some overlap with participants in the present study]; Krishnan et al., 1992, 1993), other investigators have not found such effects (e.g., Pillay et al., 1998; Lenze and Sheline, 1999). Similarly, reports of abnormalities in caudate volumes in BD have been inconsistent, with some studies reporting increased caudate volume in BD (e.g., Aylward et al., 1994; Noga et al., 2001), others reporting decreased caudate volume (e.g., Beyer et al., 2004), and still others that report no BD-associated differences in caudate volume (e.g., Dupont et al., 1995; Brambilla et al., 2001; Ong et al., 2012). Although a meta-analysis of effect sizes from separate studies of MDD and BD suggested that the caudate is smaller in MDD (Kempton et al., 2011), we did not find

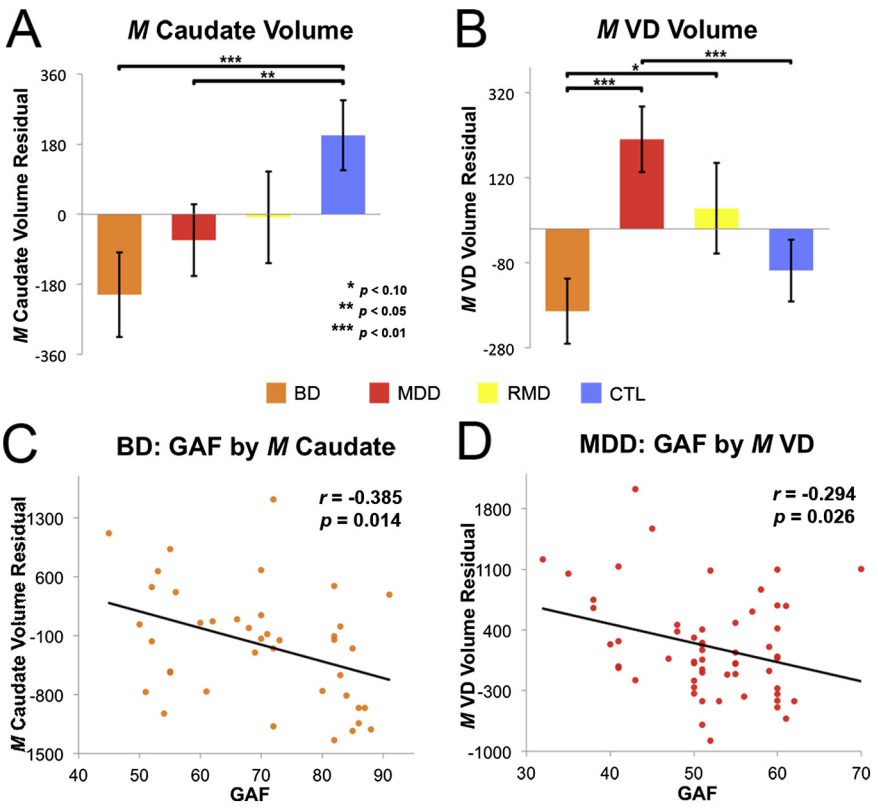

Fig. 2. Group differences in volumes and relations to global functioning. A) Caudate: greater volumes in the control (CTL) group compared to the Major Depressive Disorder (MDD; $p = 0.037$) and Bipolar Disorder (BD; $p = 0.004$) groups. B) ventral diencephalon (VD): greater volumes in the MDD group compared to the BD ($p = 0.001$) and CTL ($p = 0.004$) groups. Marginally larger VD volumes were found in the remitted depression (RMD) group than in the BD group ($p = 0.068$). Error bars indicate standard error of the mean. C) BD group scores on the global assessment of functioning (GAF) scale were negatively correlated with mean (*M*) caudate volumes; D) MDD group scores on the GAF scale were negatively correlated with *M* VD volumes.

evidence for this differentiation in comparing these two disorders directly. It is noteworthy that Dupont et al. (1995) did not find volumetric abnormalities differentiating caudate volumes in MDD, BD, and CTL individuals, but did find abnormal thalamus volumes in MDD and BD individuals (Dupont et al., 1995). These discrepancies may be due to the indirect meta-analytic method used by Kempton et al. that may be biased by the diverse neuroimaging and analytic methods of the included studies. Moreover, in the present study we included only currently euthymic BD I individuals, whereas Kempton et al. also included data from individuals diagnosed with BD II in a range of disease states (e.g., depressed, manic, euthymic); the disease state of the BD individuals in Dupont et al.'s study is not clearly stated, and their groups are also considerably smaller than those in the present study.

The VD is not typically studied in the context of mood disorders; indeed, the present study is only the second to implicate this structure in these disorders. The first study found that VD volume was a top-ranked endophenotype related to recurrent depression in a large cohort of Mexican-American individuals (Glahn et al., 2012). As segmented in the current study, the VD includes the hypothalamus (Makris et al., 2008; Glahn et al., 2012), a primary component of the hypothalamic-pituitary-adrenal (HPA) axis, which is involved in neuroendocrine and neurovegetative processes (Nestler et al., 2002). Neurons located in the paraventricular nucleus of the hypothalamus secrete corticotropin-releasing factor (CRF), which causes the anterior pituitary to release adrenocorticotropin (ACTH), which in turn causes the adrenal cortex to release the glucocorticoid cortisol (Nestler et al., 2002). Cortisol powerfully influences behavior through its action with a number of brain regions, including the hippocampus. Both MDD and BD have been hypothesized to be associated with stress-related hyperactivity of the hypothalamus (and HPA axis) through enhanced CRF transmission and/or hypercortisolism (Nestler et al., 2002; Daban et al., 2005). Our findings involving the VD suggest that abnormality of the hypothalamus is related to MDD but not to BD. Pituitary gland volumes have previously been used to study HPA-axis dysfunction (Daban et al., 2005); interestingly, researchers have found larger pituitary gland volume in BD but not in MDD individuals, relative to controls (Sassi et al., 2001). Investigators have also found higher cortisol concentration in BD than in MDD individuals during depressive episode and remission (Rybakowski and Twardowska, 1999). More generally, abnormal HPA functioning appears to more consistent during depression in BD than in MDD (see Daban et al., 2005). Taken together, BD and MDD may have unique HPA axis-related abnormalities; thus, future research should focus on delineating the roles of specific HPA-axis nodes (e.g., the hypothalamus and pituitary gland) in these disorders. In this context, Dupont et al. (1995) did not report MDD- or BD-related abnormalities in the volume of an anterior diencephalic region that included mammillary bodies, hypothalamic gray matter, and septal nucleus (Dupont et al., 1995). This discrepancy may be explained by the greater statistical power in our study than in their investigation. Thus, VD and pituitary volumes may prove useful as diagnostic markers that specifically differentiate persons with these two disorders from healthy individuals. Our ability to classify MDD vs. BD using VD and caudate volumes, and our finding of an inverse relation of GAF and VD volume within the MDD group, provides further support for the importance of VD in the pathophysiology of affective disorder.

Given the results of Kempton et al.'s meta-analysis (Kempton et al., 2011), it is noteworthy that we did not find that volumetric abnormalities of the hippocampus, pallidum, and putamen differentiated the MDD, BD, and CTL groups, nor did we find differences between the MDD and CTL groups in volume of the thalamus. These discrepancies may be due to the heterogeneity of disease states across diagnostic groups in Kempton et al.'s meta-analysis. For example, Kempton et al. included studies of euthymic, manic and depressed BD, in addition to BD I and BD II; in contrast, in the present study we included only euthymic BD I participants. Future research may be able to resolve this discrepancy by assessing subtypes of BD and MDD, or Research Domain Criteria (RDoC) constructs instead of diagnostic categories, which would permit the prediction of continuous outcomes rather than of binary classifications of disorder (Insel et al., 2010). It is also possible that heterogeneous neuroimaging or analytic methods contributed to these discrepancies. For example, the FreeSurfer automated segmentation has been posited to perform better than FSL/FIRST with respect to hippocampal segmentation (Morey et al., 2009). Relatedly, Han et al. (2006) found that field strength (i.e., 1.5 T versus 3 T), and pulse sequence parameters influence FreeSurfer's cortical thickness estimates (Han et al., 2006). Thus, this methodological heterogeneity may contribute to the discrepancies observed between our findings and previous work (Kempton et al., 2008, 2011). Furthermore, Kempton et al.'s results may have been influenced by differences in disease history across studies, whereas the disordered groups in this study did not differ from each other with respect to number of previous MDEs. Finally, although we did not find abnormalities of the whole hippocampus, given the potential role of hippocampal subfields in MDD (Small et al., 2011) and BD (Haukvik et al., 2014; Otten and Meeter, 2014), future research should benefit from comparing hippocampal subfield volumes in different forms of mood disorders.

Redlich et al. used voxel-based morphometry (VBM) to assess differences in whole-brain gray matter volume between MDD and BD individuals and found that individuals diagnosed with BD had smaller gray matter volumes than did MDD individuals in bilateral hippocampal formation, amygdala, and thalamus (Redlich et al., 2014). We did not replicate these findings in our study. It is notable, however, that Redlich et al.'s findings are also inconsistent with those reported in Kempton et al.'s meta-analysis (Kempton et al., 2011), in which MDD individuals were characterized by *smaller* volumes of the hippocampus, caudate, putamen and pallidum, and no differences were found for the thalamus. One possible explanation for the discrepancies between the current findings and Redlich et al.'s results involves the methods used in the two studies to assess gray matter volumes. Specifically, whereas Redlich et al. used VBM, we used FreeSurfer's automated segmentation. There are several important differences between FreeSurfer and VBM. For example, VBM was not designed to assess specific, *a priori* defined, subcortical structures (Makris et al., 2008; Cerasa et al., 2009), but instead, is a voxel-based procedure in which clusters may not entirely cover single regions and may also overlap multiple regions. Thus, Redlich et al. may have capitalized on the ability of VBM to assess intra-regional effects, finding effects that are spatially dispersed across subcortical structures. Our implementation of FreeSurfer was limited to *a priori* specified structures, which may have led us to obtain results that were not sufficiently consistent through an entire structure to yield significant group differences. In this context, it is noteworthy that in our study the absolute magnitude of group mean volumetric residuals of the hippocampus, amygdala, and thalamus are in the same direction as those reported by Redlich et al. Finally, we should note that whereas Redlich et al. assessed BD I individuals who were in a depressive episode, we assessed BD I individuals who were euthymic. Thus, it is possible that our findings are related to trait-like characteristics of BD, whereas Redlich et al.'s findings represent a combination of both trait- and state-like neural characteristics of BD.

The results of this study have further implications for understanding state versus trait characteristics of subcortical volumes in mood disorders. While the BD and MDD groups both exhibited reduced caudate volume compared to the CTL group, the RMD group did not differ from the CTL group. Thus, caudate volume may normalize with remission of MDD (perhaps as a marker of CSPT and/or ASPT circuit normalization) and represent a state characteristic of this disorder; in contrast, given that we found abnormal caudate volume in currently euthymic BD individuals, reduced volume in this structure may be a trait characteristic of BD. Thus, these findings offer new possibilities for differentiating MDD and BD during euthymic and remitted states. Furthermore, we found the VD to be marginally larger in the RMD than in the BD group, while there was no difference in VD volume between the RMD and MDD or CTL groups. Thus, the VD may be useful for differentiating BD individuals in a euthymic state from remitted MDD individuals. It is possible that currently manic or depressed BD individuals would exhibit abnormalities distinct from those that characterize currently depressed or remitted MDD individuals. It will be important in future research to extend the scope of assessment of abnormalities to disease states other than those examined in the current study, and to further inspect how the length of remission may influence such state-versus trait-type volumetric characteristics.

Although the individual-by-individual classifications for both MDD versus CTL and MDD versus BD groups were significant, the observed successful classifier performances (range 59.5–62.8%) were not high enough to have clinical utility. It is important to note that the machine learning paradigm used subsampling in order to maintain balance between groups. Such balancing reduces the available information for classifier training; thus, larger balanced datasets may allow for improved classification facilitated by more training examples. Additional improvement to classification may be achieved by identifying neural features that differentiate groups more effectively that do those features identified in this study, perhaps by increasing the signal-to-noise and spatial resolution of the T1-weighted MRI scan, by including features from multiple neuroimaging modalities (e.g., structural and functional MRI), or by utilizing different discriminative (e.g., logistic regression, artificial neural network) and generative (e.g., naïve Bayes, Gaussian mixture model) classification methods.

Angst et al. (2005) found that the risk of MDD converting to BD increases across the entire lifetime as the number of depressive episodes increases (Angst et al., 2005). Similarly, Goldberg and Harrow (2001) found that individuals with previous hospitalizations related to MDD converted to BD at a rate of 30% (Goldberg and Harrow, 2001). Thus, one limitation of the current study is that, even though they are all adults, some of the participants in the MDD group may go on to convert to BD. If this is the case, this would reduce the likelihood of obtaining differences in neural structure between the MDD and BD groups. Tempering this possibility, Angst et al. reported that fewer than 2% of individuals convert from MDD to BD; similarly, Goldberg and Harrow were studying individuals with histories of severe MDD. Thus, it is unlikely that the current findings are confounded by a 'mixing' of disorders by including individuals who may later convert from MDD to BD. To address this issue, future studies might assess individuals who are initially diagnosed with MDD but who later develop BD. The results of such studies will contribute to our understanding of unique and shared neural characteristics of these two disorders.

We should note four other limitations of the current study. First, the clinical groups differed in their medication use, notably in the greater use of mood stabilizers, antipsychotics, and benzodiazepines in the BD group (see Supplementary Table S2). Because neuroleptic and lithium treatment have been found to be related to increases in gray matter volume (Moore et al., 2000; Scherk and Falkai, 2006), it is unlikely that neuroleptic medications led to the observed decreased caudate volumes in the BD group (the only group with individuals who were taking antipsychotic and lithium medications). Second, the clinical groups also differed in their rates of comorbidity (Table S3). Given Lenze and Sheline's (1999) finding that non-comorbid depressed individuals did not exhibit abnormalities in caudate volume (Lenze and Sheline, 1999), it is possible that our findings are influenced by psychiatric comorbidity; future studies should address this possibility more explicitly. Third, we used a cross-sectional design in this study. Although including euthymic BD and remitted MDD participants is an important step in understanding the trait versus state nature of volumetric abnormalities in affective disorders, conducting longitudinal studies of subcortical brain volume as individuals move from currently disordered to remitted status would strengthen our conclusions in this area. Finally, the inverse relation between GAF scores and caudate volumes in the BD group was unexpected given our finding that BD was associated with smaller caudate volume compared to controls. We did not predict this relation, and it will be important, therefore, that these findings be replicated and extended with samples of MDD and BD participants.

In conclusion, we found that caudate volumes differentiated MDD and BD individuals from CTLs, and that VD volumes differentiated MDD from CTL and BD, and RMD from BD individuals. In addition, we were able to use caudate and VD volumes to classify MDD versus BD and CTL individuals. These results provide important new insights concerning subcortical volumetric differences among individuals who are experiencing different forms of mood disorders and the relations of these volumetric differences to state and trait characteristics of these disorders. Future research will benefit from improved identification of neuroanatomical features for classification, further explication of HPA-axis dysfunction in these disorders, and identification of neural markers of the transition from MDD to BD. Finally, utilizing larger datasets that include additional subtypes of affective disorder, or that use an RDoC approach, will provide important information concerning the biological foundations of disorders of mood.

Role of the funding source

This study was supported by NIMH grant MH59259 to Dr. Gotlib, a grant from the Gipuzkoako Foru Aldundia (Fellows Gipuzkoa Program) to Dr. Iglesias. The funding sources had no involvement in the study design; data collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Contributors

Mr. Sacchet conceived of the work that led to the submission. Mr. Sacchet and Dr. Gotlib designed the work that led to the submission, and drafted the manuscript. Dr. Gotlib and Ms. Livermore acquired the data. Mr. Sacchet, Ms. Livermore, and Drs. Iglesias, Glover, and Gotlib played an important role in interpreting the results, revising the manuscript, and approving the final version. No other individuals meet criteria for authorship. All authors have approved the final article.

Conflict of interest

This study was supported by a National Science Foundation Integrative Graduate Education and Research Traineeship (NSF IGERT) Recipient Award 0801700 to MDS, National Science Foundation Graduate Research Fellowship Program (NSF GRFP) DGE-1147470 to MDS, National Institute of Mental Health (NIMH) Neuroscience Research Training award T32 MH020016 to MDS, NIMH grant MH59259 to Dr. Gotlib, and a grant from the Gipuzkoako Foru Aldundia (Fellows Gipuzkoa Program) to Dr. Iglesias. The authors report nothing else to disclose.

Acknowledgments

The authors are grateful to Brian Knutson and Sheri Johnson for their contribution to data acquisition, to consultants in the Stanford Department of Statistics, Karen Larocque, and Gautam Prasad for insights related to machine learning analyses, to Anderson Winkler for help regarding data visualization, to Bob Dougherty and Tom Brosnan for their help with the assessment of scan parameters, and to David Kennedy for insights regarding neuroanatomy.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jpsychires.2015.06.002.

References

Angst J, Sellaro R, Stassen HH, Gamma A. Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions. J Affect Disord 2005;84(2–3):149–57.
Aylward EH, Roberts-Twillie JV, Barta PE, Kumar AJ, Harris GJ, Geer M, et al. Basal ganglia volumes and white matter hyperintensities in patients with bipolar disorder. Am J Psychiatry 1994;151(5):687–93.
Beyer JL, Kuchibhatla M, Payne M, Moo-Young M, Cassidy F, MacFall J, et al. Caudate volume measurement in older adults with bipolar disorder. Int J Geriat Psychiatry 2004;19(2):109–14.
Brambilla P, Harenski K, Nicoletti MA. Anatomical MRI study of basal ganglia in bipolar disorder patients. Psychiatry Res 2001;106(2):65–80.
Cerasa A, Messina D, Nicoletti G, Novellino F, Lanza P, Condino F, et al. Cerebellar atrophy in essential tremor using an automated segmentation method. AJNR Am J Neuroradiol 2009;30(6):1240–3.
Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273–97.
Daban C, Vieta E, Mackin P, Young AH. Hypothalamic-pituitary-adrenal Axis and bipolar disorder. Psychiatr Clin North Am 2005;28(2):469–80.
de Almeida JRC, Phillips ML. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biol Psychiatry 2013;73(2):111–8.
De Martino F, Valente G, Staeren N, Ashburner J, Goebel R, Formisano E. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. Neuroimage 2008;43(1):44–58.
Dupont RM, Jernigan TL, Heindel W, Butters N, Shafer K, Wilson T, et al. Magnetic resonance imaging and mood disorders: localization of White matter and other subcortical abnormalities. Arch Gen Psychiatry. Am Med Assoc 1995;52(9):747–55.
Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale: a procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry. Am Med Assoc 1976;33(6):766–71.
Fagiolini A, Kupfer DJ, Masalehdan A, Scott JA, Houck PR, Frank E. Functional impairment in the remission phase of bipolar disorder. Bipolar Disord 2005;7(3):281–5. Munksgaard International Publishers.
First MB, Dibbon M, Spitzer RL, Williams JB. Structured clinical interview for DSM-IV-TR. Washington: American Psychiatric Association; 2004.
Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci U S A 2000;97(20):11050–5.
Fischl B, Salat DH, Busa E, Albert M, Dieterich M. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341–55.
Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW, Charlesworth JC, et al. High dimensional endophenotype ranking in the search for major depression risk genes. Biol Psychiatry 2012;71(1):6–14.
Gold JI. Linking reward expectation to behavior in the basal ganglia. Trends Neurosci 2003;26(1):12–4.
Goldberg JF, Harrow M. Risk for bipolar illness in patients initially hospitalized for unipolar depression. Am J Psychiatry 2001;158:1265–70.

Hamilton JP, Etkin A, Furman DJ, Lemus MG, Johnson RF, Gotlib IH. Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural response data. Am J Psychiatry 2012;169(7):693–703.
Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. Neuroimage 2006;32(1):180–94.
Hasselbalch BJ, Knorr U, Kessing LV. Cognitive impairment in the remitted state of unipolar depressive disorder: a systematic review. J Affect Disord 2011;134(1–3):20–31.
Haukvik UK, Westlye LT, Mørch-Johnsen L, Jørgensen KN, Lange EH, Dale AM, et al. Vivo hippocampal subfield volumes in schizophrenia and bipolar disorder. Biol Psychiatry 2014;77(6):581–8.
Insel TR, Cuthbert BN, Garvey MA, Heinssen RK, Pine DS, Quinn KJ, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010;167(7):748–51.
Johnson SL, Carver CS, Gotlib IH. Elevated ambitions for fame among persons diagnosed with bipolar I disorder. J Abnorm Psychol 2012;121(3):602–9.
Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci 1998;1(5):411–6. Nature Publishing Group.
Kempton MJ, Geddes JR, Ettinger U, Williams SCR, Grasby PM. Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. Arch Gen Psychiatry 2008;65(9):1017.
Kempton MJ, Salvador Z, Munafò MR, Geddes JR, Simmons A, Frangou S, et al. Structural neuroimaging studies in major depressive disorder: meta-analysis and comparison with bipolar disorder. Arch Gen Psychiatry 2011;68(7):675–90.
Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of 12-Month DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 2005;62(6):617–27.
Kim MJ, Hamilton JP, Gotlib IH. Reduced caudate gray matter volume in women with major depressive disorder. Psychiatry Res Neuroimaging 2008;164(2):114–22.
Krishnan KRR, McDonald WM, Escalona PR, Doraiswamy PM, Na C, Husain MM, et al. Magnetic resonance imaging of the caudate nuclei in depression: preliminary observations. Arch Gen Psychiatry 1992;49(7):553–7.
Krishnan KRR, McDonald WM, Doraiswamy PM, Tupler LA, Husain M, Boyko OB, et al. Neuroanatomical substrates of depression in the elderly. Eur Arch Psychiatry Clin Neurosci 1993;243(1):41–6.
Lan MJ, Chhetry BT, Oquendo MA, Sublette ME, Sullivan G, Mann JJ, et al. Cortical thickness differences between bipolar depression and major depressive disorder. Bipolar Disord 2014;16:378–88.
Lenze EJ, Sheline YI. Absence of striatal volume differences between depressed subjects with no comorbid medical illness and matched comparison subjects. Am J Psychiatry 1999;156:1989–91.
Levens SM, Gotlib IH. Updating positive and negative stimuli in working memory in depression. J Exp Psychol Gen 2010;139(4):654–64.
Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. The brain basis of emotion: a meta-analytic review. Behav Brain Sci 2012;35(03):121–43.
Lorenzetti V, Allen NB, Fornito A, Yücel M. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. J Affect Disord 2009;117(1–2):1–17.
Makris N, Oscar-Berman M, Jaffin SK, Hodge SM, Kennedy DN, Caviness VS, et al. Decreased volume of the brain reward system in alcoholism. Biol Psychiatry 2008;64(3):192–202.
Moore GJ, Bebchuk JM, Wilds IB, Chen G, Menji HK. Lithium-induced increase in human brain grey matter. J Lancet 2000;356(9237):1241–2.
Morey RA, Petty CM, Xu Y, Hayes JP, Wagner II HR, Lewis DV, et al. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. Neuroimage 2009;45(3):855–66.
Nestler EJ, Barrot M, DiLeone RJ, Eisch AJ, Gold SJ, Monteggia LM. Neurobiology of depression. Neuron 2002;34(1):13–25.
Neumeister A, Hu X-Z, Luckenbaugh DA, Schwarz M, Nugent AC, Bonne O, et al. Differential effects of 5-HTTLPR genotypes on the behavioral and neural responses to tryptophan depletion in patients with major depression and controls. Arch Gen Psychiatry 2006;63(9):978.
Noga JT, Vladar K, Torrey EF. A volumetric magnetic resonance imaging study of monozygotic twins discordant for bipolar disorder. Psychiatry Res Neuroimaging 2001;106(1):25–34.
Ochsner KN, Silvers JA, Buhle JT. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. Ann N Y Acad Sci 2012;1251(1):E1–24.
Ong D, Walterfang M, Malhi GS, Styner M, Velakoulis D, Pantelis C. Size and shape of the caudate nucleus in individuals with bipolar affective disorder. Aust N Z J Psychiatry 2012;46(4):340–51.
Otten M, Meeter M. Hippocampal structure and function in individuals with bipolar disorder: a systematic review. J Affect Disord 2014;174:113–25.
Pillay SS, Renshaw PF, Bonello CM, Lafer B, Fava M, Yurgelun-Todd D. A quantitative magnetic resonance imaging study of caudate and lenticular nucleus gray matter volume in primary unipolar major depression: relationship to treatment response and clinical severity. Psychiatry Res 1998;84(2–3):61–74.
Price JL, Drevets WC. Neurocircuitry of mood disorders. Neuropsychopharmacology 2009;35(1):192–216.
Price JL, Drevets WC. Neural circuits underlying the pathophysiology of mood disorders. Trends Cogn Sci 2012;16(1):61–71.
Redlich R, Almeida JJR, Grotegerd D, Opel N, Kugel H, Heindel W, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression: a voxel-based Morphometry–Pattern classification approach. JAMA Psychiatry 2014;71:1222–30.
Romera I, Perez V, Menchon JM, Delgado-Cohen H, Polavieja P, Gilaberte I. Social and occupational functioning impairment in patients in partial versus complete remission of a major depressive disorder episode. A six-month prospective epidemiological study. Eur Psychiatry 2010;25(1):58–65.
Rubinsztein JS, Michael A, Paykel ES, Sahakian BJ. Cognitive impairment in remission in bipolar affective disorder. Psychol Med 2000;30(05):1025–36.
Rybakowski JK, Twardowska K. The dexamethasone/corticotropin-releasing hormone test in depression in bipolar and unipolar affective illness. J Psychiatr Res 1999;33(5):363–70.
Sassi RB, Nicoletti M, Brambilla P, Harenski K. Decreased pituitary volume in patients with bipolar disorder. Biol Psychiatry 2001;50(4):271–80.
Savitz J, Drevets WC. Bipolar and major depressive disorder: neuroimaging the developmental-degenerative divide. Neurosci Biobehav Rev 2009;33(5):699–771.
Savitz JB, Nugent AC, Bogers W, Roiser JP, Bain EE, Neumeister A, et al. Habenula volume in bipolar disorder and major depressive disorder: a high-resolution magnetic resonance imaging study. Biol Psychiatry 2011;69(4):336–43.
Scherk H, Falkai P. Effects of antipsychotics on brain structure. Curr Opin Psychiatry 2006;19(2):145–50.
Silverstone T, McPherson H, Li Q, Doyle T. Deep white matter hyperintensities in patients with bipolar depression, unipolar depression and age-matched control subjects. Bipolar Disord 2003;5:53–7.
Singh T, Rajput M. Misdiagnosis of bipolar disorder. Psychiatry (Edgmont) 2006;3(10):57–63.
Small SA, Schobel SA, Buxton RB, Witter MP, Barnes CA. A pathophysiological framework of hippocampal dysfunction in ageing and disease. Nat Rev Neurosci 2011;12:585–601.
Versace A, Almeida JRC, Quevedo K, Thompson WK, Terwilliger RA, Hassel S, et al. Right orbitofrontal corticolimbic and left corticocortical White matter connectivity differentiate bipolar and unipolar depression. Biol Psychiatry 2010;68(6):560–7.
Victor SE, Johnson SL, Gotlib IH. Quality of life and impulsivity in bipolar disorder. Bipolar Disord 2011;13(3):303–9.
World health organization depression fact sheet No 369 [Internet]. 2012 [cited 2014 Aug 13]. Available from, http://www.who.int/mediacentre/factsheets/fs369/en/.
Yildiz A, Sachs GS, Dorer DJ, Renshaw PF. 31 P nuclear magnetic resonance spectroscopy findings in bipolar illness: a meta-analysis. Psychiatry Res 2001;106(3):181–91.