Journal of Affective Disorders 281 (2021) 303–311

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

Neuroimaging correlates of emotional response-inhibition discriminate between young depressed adults with and without sub-threshold bipolar symptoms (Emotional Response-inhibition in Young Depressed Adults)

Jungwon Cha [1], Sidra Speaker [2], Bo Hu [3], Murat Altinay [1], Parashar Koirala [1], Harish Karne [1], Jeffrey Spielberg [4], Amy Kuceyeski [5], Elvisha Dhamala [5], Amit Anand [1,*]

[1] Center for Behavioral Health, Cleveland Clinic
[2] Cleveland Clinic Lerner College of Medicine
[3] Center for Quantitative Health Sciences, Cleveland Clinic
[4] Department of Psychiatry, University of Delaware
[5] Department of Radiology, Weill Cornell Medicine, Brain and Mind Research Institute, Weill Cornell Medicine

ARTICLE INFO

Keywords:
Bipolar Depression
fMRI
Emotional response inhibition
Subthreshold bipolar disorder Young adults
Machine learning

ABSTRACT

Background: Many subjects with major depression (MDD) exhibit subthreshold mania symptoms (MDD+). This study investigated, for the first time, using emotional inhibition tasks, whether the neural organization of MDD+ subjects is more similar to bipolar depression (BDD) or to MDD subjects without subthreshold bipolar symptoms (MDD-).

Method: This study included 118 medication-free young adults (15 – 30 yrs.): 20 BDD, 28 MDD+, 41 MDD- and 29 HC subjects. Participants underwent fMRI during emotional and non-emotional Go/No-go tasks during which they responded for Go stimuli and inhibited response for happy, fear, and non-emotional (gender) faces No-go stimuli. Univariate linear mixed-effects (LME) analysis for group effects and multivariate Gaussian Process Classifier (GPC) analyses were conducted.

Results: MDD- group compared to both the BDD and MDD+ groups, exhibited significantly lower activation in parietal, temporal and frontal regions (cluster-wise corrected $p < 0.05$) for emotional inhibition conditions vs. non-emotional condition. GPC classification of emotional (happy + fear) vs. non-emotional response-inhibition activation pattern showed good discrimination between BDD and MDD- subjects (AUC: 0.70; balanced accuracy: 70% (corrected $p = 0.018$)) as well as between MDD+ and MDD- subjects (AUC: 0.72; balanced accuracy: 67% (corrected $p = 0.045$)) but less efficient discrimination between BDD and MDD+ groups (AUC: 0.68; balanced accuracy: 61% (corrected $p = 0.273$)). Notably, classification of the MDD- group was weighted for left amygdala activation pattern.

Limitations: Results also need to be tested in a different independent dataset.

Conclusion: Using an fMRI emotional Go-Nogo task, MDD- subjects can be discriminated from BDD and MDD+ subjects.

Introduction

It has been estimated that 30-55% of patients presenting with major depression (MDD) have additional sub-threshold (hypo) mania symptoms (MDD+) per DSM-IV-TR criteria. MDD+ subjects are thought to be at a higher risk for developing bipolar disorder (BD) (Coryell et al., 1995; Fiedorowicz et al., 2011) compared to MDD subjects who do not have subthreshold mania symptoms (MDD-). However, the neurobiology of MDD+ subjects has not been adequately studied to determine whether MDD+ subjects are more akin to BDD or to MDD-subjects. In particular, young subjects presenting with depression and subthreshold mania (MDD+) present a conundrum for clinicians, as the duration of illness is usually too short to determine the diagnosis's stability. Therefore, there is a critical need to differentiate and identify objective classifiers, which may help in differentiating BDD, MDD- and MDD+ in young adult subjects.

* Corresponding author at: Center for Behavioral Health, 9500 Euclid Avenue P57, Cleveland, OH 44195.
E-mail address: ananda@ccf.org (A. Anand).

https://doi.org/10.1016/j.jad.2020.12.037
Received 29 May 2020; Received in revised form 25 September 2020; Accepted 7 December 2020
Available online 10 December 2020

Neuroimaging characteristics for BD may provide such potential classifiers. Neuroimaging studies in BD have implicated abnormal corticolimbic activation and connectivity in BD though no single abnormality has been robustly validated through replication studies (Phillips & Swartz, 2014; Strakowski et al., 2012). fMRI task-induced activation paradigms involving affective faces have frequently been used in studies of mood disorders, including studies comparing MDD and BDD in order to isolate BDD-specific characteristics (Han, De Berardis, Fornaro, & Kim, 2019; Redlich et al., 2014; Redlich et al., 2015). However, a critical gap remains, given that very few studies have examined this question *within* the MDD group; in other words, comparing MDD+ to MDD- subjects. Fournier and colleagues reported that greater right-amygdala activity to the happy faces in adults with MDD was associated with higher levels of subthreshold manic symptoms experienced across the lifespan (Fournier et al., 2013). A recent study of remitted young depressed subjects found a difference in resting state brain connectivity in subjects with subthreshold bipolar symptoms (Kling et al., 2018). Thus, there are indications that affect-face-related activation may be used to characterize MDD+ individuals, but this has not been explicitly tested. In the present study, we have attempted to fill this critical gap in the literature by testing whether fMRI activation during response-inhibition to emotional tasks can differentiate and accurately classify MDD+ and MDD- individuals.

BD is frequently associated with difficulty inhibiting responses, particularly to emotion-laden stimuli. One commonly used task to assay these difficulties is a variant of the Go/No-go task, in which a subject has to inhibit responses to a rare No-go stimulus in the context of speeded responses to a common Go stimulus. Both non-emotional (e.g., letters) and emotional (e.g., faces depicting an emotion) stimuli are commonly used in this paradigm.

A number of studies have used emotional Go/No-go tasks to investigate BD-related activation differences. For example, BD has been associated with differential activation on emotional Go/No-go tasks in the temporal cortex, OFC, insula, caudate, and the dorsal anterior and posterior cingulate cortices (Hummer et al., 2013; Townsend et al., 2012; Wessa et al., 2007).

Although several studies have used close relatives of bipolar subjects to identify high-risk endophenotypes (Frangou, Dima, & Jogia, 2017; Piguet, Fodoulian, Aubry, Vuilleumier, & Houenou, 2015; Vierck, Porter, & Joyce, 2015), MDD+ subjects have not yet been used to identify clinically applicable activation patterns, with the ultimate goal of applying these patterns in clinical work to identify at-risk individuals.

Typical fMRI studies in this area employ univariate analyses to find group-related differences. Though information gained via these methods provides valuable insights into possible BDD pathophysiology, they have less individual-level-clinical utility for differentiating between groups (Amit Etkin, 2019; Frangou et al., 2017). In comparison, multivariate machine learning techniques, such as support vector machines (SVM) or Gaussian Process Classifiers (GPC), can be used to provide individual-level classification accuracy of an fMRI task. For example, Frangou and colleagues used GPC to classify BD and relatives of BD at-risk for developing BD, using a 3-back working memory task (Frangou et al., 2017). Such individual level analyses could have utility in clinical situations which require correctly e.g., differentiating between MDD+ and MDD-.

This study aimed to ascertain the differentiating and classification-accuracy of emotional Go/No-go task-induced fMRI activation in MDD+ and MDD- subjects. Based on previously described findings, we hypothesized that during emotional response-inhibition, MDD+ and BDD participants would show a more similar pattern of activation for emotional vs. non-emotional response-inhibition tasks, which will be different from the pattern seen in MDD- subjects. To test this hypothesis, we examined, for the first time, fMRI task-induced activation differences between medication-free young subjects with BDD, MDD+, MDD-, and HC as participants completed emotional (happy, fearful) and non-emotional (gender recognition) Go/No-go tasks. Next, we used a GPC machine learning algorithm to evaluate classification performances of fMRI activation patterns and identify brain areas that contributed the most to the classification.

Methods and Materials

Subjects

Bipolar and MDD participants ages 15-30 years who were medication-free for at least 2 weeks were recruited from the outpatient psychiatry clinic at the Cleveland Clinic and by advertisement. Healthy control participants were recruited by advertisement to the community. Inclusion and Exclusion criteria are presented in the supplement.

Depression subgroup ascertainment using best practices: Three psychiatrists independently reviewed all information available for each subject and classified all MDD subjects as MDD+ or MDD- at the conclusion of the study (Koirala et al., 2019). We defined MDD+ group consisting of subjects who exhibited mania symptoms that did not satisfy full DSM-IV-TR criteria for hypomania or mania. Based on a review of previous studies (Angst et al., 2003; Fiedorowicz et al., 2011; Koirala et al., 2019; Merikangas et al., 2011; Zimmermann et al., 2009), which have used varying definitions for subthreshold mania symptoms, we formulated conservative criteria. Thus subthreshold symptoms were defined as: euphoric mood with at least 2 mania symptoms or increased irritability with 3 mania symptoms if only the latter was present, as well as if more mania symptoms were present then duration to be for less than 4 days. All Other MDD subjects who did not have subthreshold symptoms, no family history of BD, and no history of psychosis (Angst et al., 2003; Fiedorowicz et al., 2011; Koirala et al., 2019; Merikangas et al., 2011; Zimmermann et al., 2009), were ascertained to be MDD-. Subsequently, the classification of each of the subjects was discussed between all three psychiatrists and a consensus best-estimate classification was agreed upon (Nurnberger et al., 2011).

Naturalistic follow-up for change in diagnosis: as an exploratory aim of the study, MDD subjects who wanted to be treated were started on a selective serotonin reuptake inhibitor (SSRI) and followed up for up to 2 years. The behavioral course of the larger sample, which was clinically followed up, has been reported previously (Koirala et al., 2019). Any change in diagnosis to BD was recorded.

Procedures

Tasks. We used a modified emotion Go/No-go fMRI paradigm similar to that used in other studies (Hummer et al., 2013; Wessa et al., 2007) with fearful, happy and neutral facial stimuli showing both male and female faces as depicted and described in Supplemental Figure S-1. Emotional inhibition task conditions consisted of emotional No-go blocks in which (Happy or Fearful) No-go stimuli were interspersed among Neutral Go stimuli. For emotional inhibition, regardless of valence, Emotional Inhibition (Happy No-go + Fearful No-go) condition was created. For emotional inhibition tasks, a contrast with each of the emotional No-go conditions with the neutral Go condition was created. Neutral Go rather than emotional Go was used as the control stimulus for all emotional No-go blocks to avoid responses to strong repetitive emotional stimulus dominating the activation response. No-go stimuli were interspersed among Neutral (Female faces and Male faces) Go stimuli (Supplemental Figure S-1). Non-emotional inhibition condition was constructed consisting of (Female + Male) No-go blocks and Go blocks.

Behavioral Analysis. To examine performance differences between groups, Go and No-go accuracies and reaction times were compared using SPSS (Version 21, IBM, Chicago, Illinois) via one-way analysis of variance (ANOVA) tests with post hoc t-test and Bonferroni correction for multiple comparisons as shown in Supplemental Table 1.

Imaging procedure and preprocessing: details are presented in the online supplement Functional neuroimaging data analysis:

A General Linear Model was implemented with a delayed boxcar waveform to model blood oxygen level-dependent (BOLD) signal changes in relation to conditions. The data was scaled at the first level to deal with the arbitrariness of BOLD signal varying across brain regions and subjects (within-subject) (Gang Chen, Taylor, & Cox, 2017). Six movement parameters, calculated during the realignment, were used as covariates of no interest to account for variations in the signal from movement artifacts. First-level analyses produced beta images for each condition of interest (happy no-go, fearful no-go, neutral go, male no-go, female no-go, male go, and female go) from each participant using 3dDeconvolve from AFNI.

A mask was created of the average effect of each task condition across all subjects at voxel-wise threshold $p < .001$ (uncorrected), representing the cluster-wise corrected significance of $p < .05$ to be used for second-level group effect analysis for each condition respectively. To calculate the cluster-wise corrected significance threshold for a given voxel-wise threshold p-value, we used 3dClustsim's non-Gaussian autocorrelation function (ACF), which uses the smoothness parameters estimated by the 3dFWHMx function from AFNI (Cox, Chen, Glen, Reynolds, & Taylor, 2017; Forman et al., 1995).

We conducted a second-level linear mixed-effects (LME) analysis with a whole-brain mask using 3dLME function in AFNI (G. Chen, Saad, Britton, Pine, & Cox, 2013). F-statistics for the main effect of group x condition were calculated for these contrasts using 3dLME. We used group as a between-subject factor and condition as a within-subject factor, controlling for task accuracy and reaction time with a within-subject quantitative variable, and age, race, and scanner type with between-subject covariates to identify the effect of group on the pre-specified contrasts.

Additionally, we conducted main effects of group x condition analyses of Go only blocks contrasted with Rest condition, to ensure that Go blocks were not driving significant results in No-go – Go contrasts. Significant Go/No-go results are only reported only if analogous Go vs. Rest clusters were not independently significant.

For differences across patient groups, we performed a second-level linear mixed-effects (LME) analysis for four groups using age, race, scanner type, reaction time, and task accuracy as covariates to exclude the effect of these confounding variables.

The main effect of group x condition was examined at voxel-wise threshold $p < .001$ (uncorrected) threshold and cluster-wise corrected significance of $p < .01$ using the average effect of condition mask. We used a more stringent cluster-wise corrected threshold of p-value to include the most significant clusters. The location of the cluster was identified using the location of the peak voxel using the MNI Atlas. Cluster sizes required for corrected significance are given below under Results.

Subject-specific beta coefficients from each main effect cluster were analyzed in SPSS (Version 21, IBM, Chicago, Illinois) for significant post-hoc pairwise group differences using Bonferroni correction for multiple comparisons (Table 1).

Multivariate pattern classification

Binary Gaussian Process Classifier (GPC) was performed in the Pattern Recognition for Neuroimaging Toolbox (PRoNTo) (www.mlnl.cs.ucl.ac.uk/pronto/) using contrasts based on whole-brain beta images for each condition of interest (happy No-go, fearful No-go, emotional No-go, neutral go, male No-go, female No-go, non-emotional No-go, male go, and female go).

GPC is a supervised learning method for classification, which builds a model with (Gaussian) predictive probabilities of class membership, which allows us to estimate uncertainty in the prediction (Rasmussen, 2003). Age, race, scanner type, and 17-item Hamilton Depression (HAM-D) were used as covariates in the classifiers. Each classifier used the mask computed by the average effect of each condition across all subjects in the corresponding group-pair at voxel-wise threshold of $p <$

Table 1
Significant Results of Main Effects for Group x Condition and Post-hoc Comparisons with Bonferroni Correction for Multiple Comparisons.

Main Effects for Group x Condition				**Post-hoc Comparisons**
Activation Area	**Cluster Size**	**Peak MNI (X Y Z)**	**Peak F with Degree of Freedom (3, 111)**	**Post-hoc test Bonferroni Corrected p < 0.05**
Emotional Face Inhibition				
Right Cuneus	133	20, -79, 43	8.727	BDD > MDD-, MDD+ > MDD-
Non-Emotional Face Inhibition				
Right Temporal Pole	128	35, 23, -27	10.515	HC < MDD-
Happy Face Inhibition vs. Non-Emotional Face Inhibition				
Right Angular Gyrus	214	45, -69, 40	8.324	BDD > MDD-, MDD+ > MDD-, HC > MDD-
Right Medial Temporal Pole	205	58, 11, -30	13.836	HC > MDD-, MDD+ > MDD-
Right Precuneus	188	2, -67, 56	8.221	BDD > MDD-, HC > MDD-
Right Middle Frontal Gyrus	112	35, 31, 48	8.711	BDD > MDD-,
			8	HC > MDD-
Right Temporal Pole	79	30, 11, -24	7.978	HC > MDD-
Left Inferior Temporal Gyrus	66	-62, -49, -17	12.646	HC > MDD-, MDD+ > MDD-
Fearful Face Inhibition vs. Non-Emotional Face Inhibition				
Right Medial Temporal Pole	55	58, 8, -30	9.184	MDD+ > MDD-
Emotional Face Inhibition vs. Non-Emotional Face Inhibition				
Right Posterior Cingulate Cortex	209	10, -39, 8	9.566	BDD > MDD-, HC > MDD-
Right Precuneus	181	8, -72, 53	8.964	BDD > MDD-, HC > MDD-
RightMedial Temporal Pole	136	58, 11, -30	13.076	BDD > MDD-, MDD+ > MDD-
Right Angular Gyrus	121	40, -57, 38	7.788	BDD > MDD-
			7	HC > MDD-
Right Hippocampus (within 4mm)	84	40, -4, -20	9.756	MDD+ > MDD-
Right Superior Frontal Gyrus	69	28, 68, 3	11.089	HC > MDD-
Right Fusiform Gyrus	64	30, 13, -44	10.340	none
Right Fusiform Gyrus	58	35, -29, -22	7.994	BDD > MDD-, MDD+ > MDD-
Right Superior Temporal Gyrus	54	58, -24, 3	7.468	HC > MDD-
Right Cerebellum (IX)	45	10, -47, -47	8.593	BDD > MDD-, HC > MDD-
Left Inferior Temporal Gyrus	41	-62, -49, -17	11.274	HC > MDD-, MDD+ > MDD-

.001 (uncorrected), cluster-wise corrected significance of $p < .05$. Each classifier was trained using a leave-one-out per group cross-validation (CV). In order to perform the CV, one subject in each group was selected and then allocated to the test set. The primary metric for evaluation of classification performance was the area under the receiver operating characteristics (ROC) curve (AUC). Moreover, due to an imbalanced dataset, standard accuracy may not be a valid way to examine the accuracy of the classifier. In contrast, sensitivity (true positive rate) and specificity (true negative rate) better accommodate imbalanced data. Thus we calculated - balanced accuracy (average of sensitivity and specificity), which is thought to be a more valid metric for interpreting the classification for an imbalanced dataset (Brodersen et al., 2011). A threshold of 0.5 for class-labels was used. Next, permutation testing was performed for each classifier with 1000 permutations, which provided a p-value for the balanced accuracy of classification (Golland & Fischl, 2003; Nichols & Holmes, 2002; Noirhomme et al., 2014).

Finally, we computed whole-brain discrimination maps for each classifier to visualize each voxel's quantitative contribution to the classifier's decision.

Classification between pairs of diagnostic groups BDD vs. MDD-, BDD vs. MDD+, and MDD+ vs. MDD- were analyzed. Separately, we explored whether the addition of MDD+ to BDD or addition to MDD- led to better discrimination. This was done by looking at the classification between (BDD + MDD+) vs. MDD- group and the BDD vs. (MDD- + MDD+).

Results

Demographics

One-hundred seventy-seven subjects were enrolled in the study. Thirty participants were excluded due to excessive motion or problematic quality during image acquisition. Five participants were excluded because they only had a family history of BD (4 subjects) or psychosis (1 subject) only but no subthreshold symptoms. These subjects were excluded to reduce heterogeneity in the MDD+ sample. Eight HCs were also excluded in order to match race across groups. Two participants from the HC group, later on, were discovered to have had a family history of mental illness and were excluded. One BDD participant was excluded who was judged to give unreliable information. Three participants were excluded since they were on hormones and transition to the opposite gender. The final analyses included 118 medication-free subjects: 20 BDD (8 bipolar I and 12 bipolar II), 28 MDD+, 41 MDD-, and 29 HC subjects. The BDD subtypes were combined as a whole and compared with the MDD+ and MDD- groups as there were not a sufficient number of subjects to conduct analysis taking account of the BD I and II subtypes.

Demographic Characteristics of each group are presented in Table 2.

Three patient group analysis: All demographic factors significantly different between the 4 groups and all illness characteristics significantly different between the 3 patient groups were used as covariates in the analyses. The MDD-group had a lower mean HAM-D score than the MDD+ and BDD groups. Therefore, to examine for any effect of differences in HAM-D depression scores on differences seen between the three patient groups results seen in the four-group analysis, we also conducted a three-group analysis for the patient groups while excluding the HC group. In this three-group analysis, the HAM-D score was additionally used as a covariate beside the other covariates included in the four group analyses. YMRS scores were low in all depressed patient groups but were significantly different in the BDD and MDD+ groups compared to the MDD- group. However, YMRS scores were not used as covariates as they were the inherent characteristics of the MDD+ and the BDD groups. The results of the three-group analysis were mostly the same as the four-group analysis in terms of brain regions and post-hoc differences between the patient groups (Supplementary Table 2). The four-group analysis, which includes healthy subjects in the LME model, is presented below.

Task Performance

After correction for multiple pairwise group differences, no significant differences were seen between groups for any of the Go or No-Go conditions for accuracy and reaction times (Supplemental Table 1). The only exception was for the Fearful No-go conditions for which the healthy control group showed greater accuracy than the BDD group. However, to further make sure that accuracy or reaction times were not driving any of the imaging he results we used them both as covariates in all second-level imaging analyses.

Table 2
Demographics and Illness Characteristics.

Demographics four group ANOVA		**BDD (N=20)**	**MDD+ (N=28)**	**MDD- (N=41)**	**HC (N=29)**	**3-Group p-value**	**4-Group p-value**
Age (years)(mean (SD))		24.8(4.1)	23.1(3.5)	25.1(3.8)	24.8(3.4)	0.083	0.121
Female (n (%))		14(70%)	24(86%)	27(66%)	18(62%)	0.178	0.211
Race	Caucasian (n (%))	17(85%)	21(75%)	39(95%)	26(90%)		
	African American (n (%))	3(15%)	7(25%)	2(5%)	3(10%)	0.054	0.097
Scanner	Scanner 1 (n (%))	4(20%)	16(57%)	11(27%)	7(24%)		
	Scanner 2 (n (%))	16(80%)	12(43%)	30(73%)	22(76%)	0.010	0.013
Handedness (n of right handed (%))		18(90%)	25(86%)	38(93%)	22(88) ¶	0.641	0.818
Illness characteristics three group ANOVA							
HAM-D 17 item (mean (SD))		18.6(5.1)	19.7(3.7)	16.6(3)		0.004	
YMRS (mean (SD))		2.1(2.5)	2.5(2.8)	0.8(1.2)		0.002	
Age at First Episode (years)(mean (SD))		13.3(3.7)	14.7(4.3)	15.3(4.4)		0.201	
Duration of illness (years)(mean (SD))		10.5(5.6)	8.4(5.6)	9.8(5.4)		0.390	
Medication Free Period (weeks)(mean (SD))*		108.5(108.9)	70(71.6)	95.7(118.3)		0.551	
Classification Characteristics							
Bipolar I(n): Bipolar II(n)		8:12					
First degree family history (n (%))		6(30%)	4(14%)	0(0)			
Psychosis (n (%))		2(10%)	3(11%)	0(0)			

¶ 4 subjects are missing information.
* There were no significant differences in the distribution of medication naïve participants across groups.
ANOVA test was performed on the continuous variables and the chi-squared test was performed on the categorical variables.
BDD: bipolar depressed group
MDD+: major depressive disorder patients with subthreshold mania symptoms
MDD-: major depressive disorder patients who no subthreshold mania symptoms, history of psychosis or family history of bipolar disorder
HC: Healthy Control

Condition Effects

The effect of each of the following conditions: Happy face inhibition (Happy Nogo – Neutral Go), Fearful face inhibition (Fearful Nogo-Neutral Go), Emotional face inhibition ((Happy or Fear) Nogo – Neutral Go) and Non-emotional face inhibition ((Male or Female) Nogo – (Female or Male) Go) are presented in Supplementary figure S-2 and Supplementary Table 3.

Average effect of condition was examined at voxel-wise threshold of $p < .001$ (uncorrected) and masks at cluster-wise corrected significance of $p < .05$ (calculated using 3dClustsim's auto-correlation function (ACF)) were made at $k \geq 86$ voxels (happy face inhibition), 66 voxels (fearful face inhibition), 83 voxels (emotional face inhibition), 95 voxels (non-emotional face inhibition), 102 voxels (happy face inhibition vs. non-emotional face inhibition), 75 voxels (fearful face inhibition vs. non-emotional face inhibition), and 86 voxels (emotional face inhibition vs. non-emotional face inhibition). The average effects of condition masks were used for group analysis as described below.

Within-Group Imaging Results

We conducted one-sample t-tests for each relevant group and visualized differentially activated clusters masked for average effects of the condition in each of the groups for each of the conditions. AFNI's 3dClustSim with ACF parameters was used to estimate the cluster size threshold for clusters at voxel-wise threshold of $p < .005$ (uncorrected) for the MDD+, MDD- and HC groups which would represent the cluster-wise corrected significance of $p < 0.05$. Due to the smaller size of the BDD group, the voxel-wise threshold was increased to $p < .01$ for graphical purposes (Supplemental Figures S-3 to S-9). Within-group imaging results are detailed in the Supplemental Material.

Next, we conducted a linear mixed effects univariate analysis to identify between group differences.

Between-Group Imaging Results

For main effects for group x condition, AFNI's 3dClustSim with ACF parameters was used to estimate the cluster size threshold for clusters at voxel-wise threshold of $p < .001$ (uncorrected) which would represent the cluster-wise corrected significance of $p < .01$. 3dClustSim estimated significant cluster size thresholds of $k = 21$ voxels (happy face inhibition), 60 voxels (fearful face inhibition), 70 voxels (emotional face inhibition), 57 voxels (non-emotional face inhibition), 61 voxels (happy face inhibition vs. non-emotional face inhibition), 42 voxels (fearful face inhibition vs. non-emotional face inhibition), and 39 voxels (emotional face inhibition vs. non-emotional face inhibition). Significant results are depicted in Table 1. Also, presented are post-hoc pairwise group comparisons at $p < .05$ corrected for multiple comparisons.

Happy Face Inhibition: There were no significant main effects for group x condition during happy face inhibition.

Fearful Face Inhibition: There were no significant main effects for group x condition during fearful face inhibition.

Emotional Face Inhibition: There was a significant main effect for group x condition in right cuneus, with post-hoc analysis showing lower activation in the MDD- group compared to the MDD+ and BDD groups (Supplemental Figure S-10(a)).

Non-Emotional Face Inhibition: There was a significant main effect for group x condition in the right temporal pole, which had lower activation in the HC group compared to the MDD- but showed no difference with the BDD group and the MDD+ group (Supplemental Figure S-10(b)).

The contrast between emotional conditions vs. the non-emotional condition revealed more extensive differences between groups.

Happy Face Inhibition vs. Non-Emotional Face Inhibition: There was a significant main effect for group x condition in the right angular gyrus, right medial temporal pole, right precuneus, right middle frontal gyrus, right temporal pole, and left inferior temporal gyrus as shown in Supplemental Figure S-11. In these regions, the left inferior temporal gyrus showed lower activation in the MDD- group compared to the MDD+ and HC groups. Moreover, the right angular gyrus and the right precuneus showed lower activation in the MDD- group compared to the BDD and HC groups.

Fearful Face Inhibition vs. Non-Emotional Face Inhibition: There was a significant main effect for group x condition in the right medial temporal pole with post-hoc analyses showing lower activation in the MDD- group compared to the BDD and MDD+ groups as shown in Supplemental Figure S-12.

Emotional Face Inhibition vs. Non-Emotional Face Inhibition: This comparison showed more extensive differences between the groups. There were significant main effects for group x condition in the right posterior cingulate cortex, right precuneus, right medial temporal pole, right angular gyrus, right hippocampus (within 4mm), right superior frontal gyrus, right fusiform gyrus, right superior temporal gyrus, right cerebellum (IX), and left inferior temporal gyrus as shown in Fig. 1. Most of these significant regions showed lower activation in the MDD- group compared to the BDD and MDD+ groups.

Bipolar I vs. Bipolar II: as some studies have reported differences in functional imaging parameters between bipolar I and II subtypes (Cardoso de Almeida & Phillips, 2013), we conducted a preliminary analysis of the differences between the two groups in the post-hoc analysis. No corrected significant results were found between the two groups.

Next multivariate machine learning was applied to investigate the individual-level classification accuracy of the Go-No-go task for subjects in each group (Frangou et al., 2017).

Classification Results

Classification of beta images using GPC was performed only on the contrast which showed significant main effects for group x condition at cluster-wise corrected significance $p < .01$ as described above. Classification results are detailed in Table 3. Receiver Operating Characteristics (ROC) curves with Area Under Curve (AUC) of classification are depicted from Fig. 2 and Supplementary Figure S-13 and S-14, along with discrimination weight maps. We used Bonferroni correction to account for the multiple pairwise group comparisons. Results were considered significant if the corrected p-value for the balanced accuracy of classification $< .05$.

Happy Face Inhibition vs. Non-Emotional Face Inhibition: Classifier using beta images for the happy face inhibition vs. non-emotional inhibition contrast showed significant discrimination (AUC: 0.73; balanced accuracy: 70% (corrected $p = 0.015$)) between BDD vs. MDD- subjects. The largest discriminating clusters were in temporal and parietal regions. The classifier between (BDD & MDD+) and MDD- groups also presented significant discrimination (AUC: 0.70; balanced accuracy: 66% (corrected $p = 0.01$)). Discrimination maps showed that the temporal and parietal brain regions contributed most to these classifiers as shown in Supplemental Figure S-13.

Fearful Face Inhibition vs. Non-Emotional Face Inhibition: There was no significant discrimination for the fearful face inhibition vs. non-emotional inhibition contrast. The classifier between BDD and MDD- groups presented trend level significant discrimination (AUC: 0.66; balanced accuracy: 56% (corrected $p = 0.054$)). Discrimination maps showed that the temporal and parietal brain regions contributed most to these classifiers as shown in Supplemental Figure S-14.

Emotional Face Inhibition vs. Non-Emotional Face Inhibition: Emotional inhibition vs. non-emotional inhibition contrast (AUC: 0.70; balanced accuracy: 70% (corrected $p = 0.018$)) showed significant discrimination between BDD and MDD-. The classifier between MDD+ and MDD- groups provided the best discrimination (AUC: 0.72; balanced accuracy: 67% (corrected $p = 0.045$)). The emotional inhibition vs. non-emotional inhibition contrast (AUC: 0.71; balanced accuracy: 65% (corrected $p = 0.024$)) also showed significant discrimination between

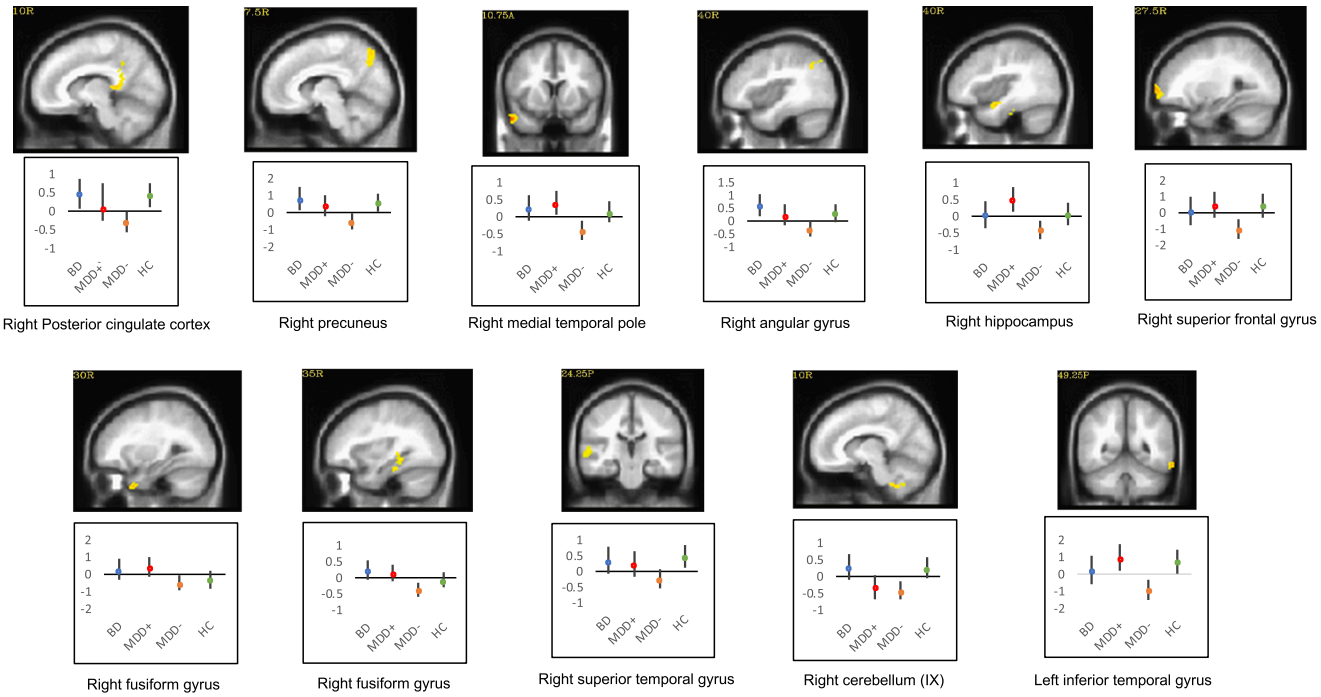


Fig. 1. Group activation differences during emotional face inhibition vs. non-emotional face inhibition. The figures are shown at cluster size k ≥ 39 voxels, cluster-wise corrected significance of p < .01. Mean and 95% confidence intervals for average activity within main effect cluster are represented. There were significant main effects for group x condition in 11 ROIs.

Table 3
Classification Results (A) for the BDD and MDD- groups, the BDD and MDD+ groups and the MDD+ and MDD- groups, and (B) for the (BDD & MDD+) and MDD- groups, and the BDD and (MDD+ & MDD-) groups Based on Beta Images Masked by Average Effect of Condition Mask Using Gaussian Process Classifier.

A.

	BDD vs. MDD-			BDD vs. MDD+			MDD+ vs. MDD-		
	Happy Face Inhibition vs. Non-Emotional Face Inhibition	Fearful Face Inhibition vs. Non-Emotional Face Inhibition	Emotional Face Inhibition vs. Non-Emotional Face Inhibition	Happy Face Inhibition vs. Non-Emotional Face Inhibition	Fearful Face Inhibition vs. Non-Emotional Face Inhibition	Emotional Face Inhibition vs. Non-Emotional Face Inhibition	Happy Face Inhibition vs. Non-Emotional Face Inhibition	Fearful Face Inhibition vs. Non-Emotional Face Inhibition	Emotional Face Inhibition vs. Non-Emotional Face Inhibition
AUC	**0.73**	**0.66**	**0.70**	**0.66**	**0.69**	**0.68**	**0.62**	**0.66**	**0.72**
Accuracy	74	67	75	58	67	63	62	70	70
Balanced Acc.	70	56	70	56	66	61	60	65	67
Sensitivity	60	25	55	45	60	55	50	39	54
Specificity	81	88	85	68	71	68	71	90	80
PPV	60	50	65	50	60	55	54	73	65
NPV	81	71	80	63	71	68	67	69	72
Corrected p-value	**0.015**	**0.054**	**0.018**	**0.711**	**0.099**	**0.273**	**0.234**	**0.078**	**0.045**

B.

	(BDD & MDD+) vs. MDD-			BDD vs. (MDD+ & MDD-)		
	Happy Face Inhibition vs. Non-Emotional Face Inhibition	Fearful Face Inhibition vs. Non-Emotional Face Inhibition	Emotional Face Inhibition vs. Non-Emotional Face Inhibition	Happy Face Inhibition vs. Non-Emotional Face Inhibition	Fearful Face Inhibition vs. Non-Emotional Face Inhibition	Emotional Face Inhibition vs. Non-Emotional Face Inhibition
AUC	**0.70**	**0.63**	**0.71**	**0.65**	**0.66**	**0.53**
Accuracy	66	57	65	74	75	69
Balanced Acc.	66	58	65	58	59	55
Sensitivity	65	50	65	30	30	30
Specificity	68	66	66	87	88	80
PPV	70	63	69	40	43	30
NPV	62	53	61	81	81	80
Corrected p-value	**0.01**	**0.228**	**0.024**	**0.152**	**0.082**	**0.344**

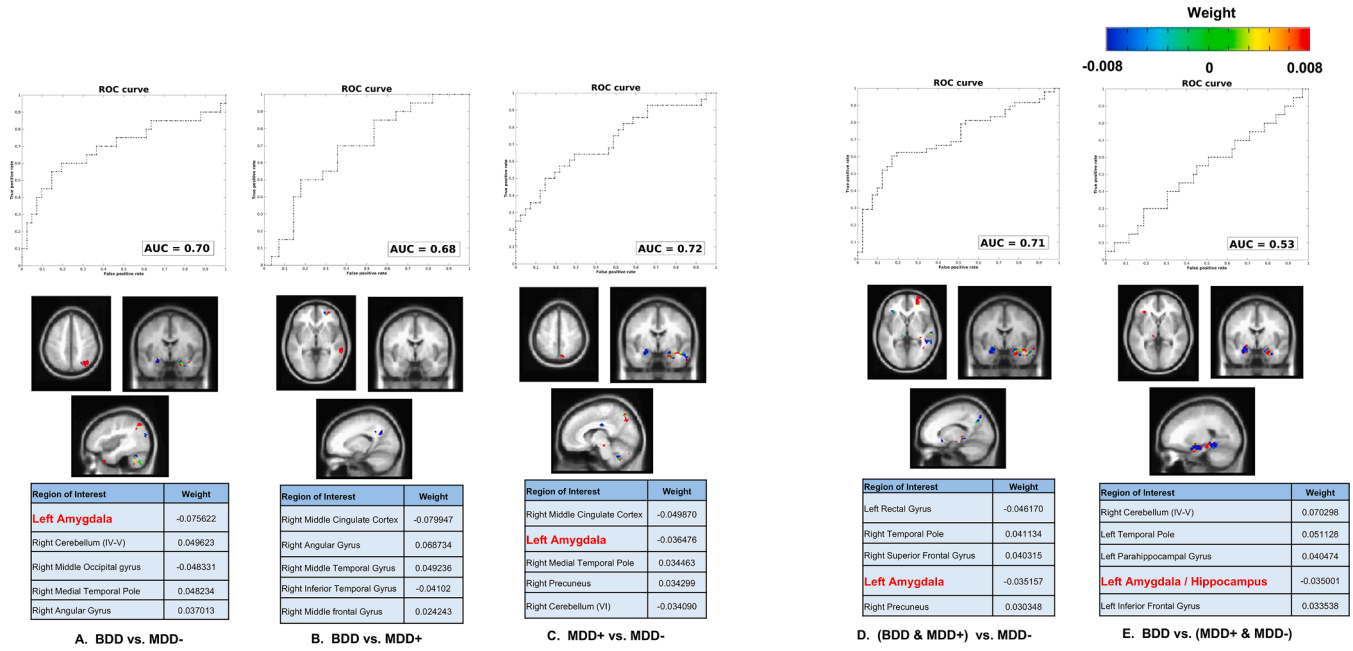
BDD, bipolar depressed group; MDD+, major depressive disorder patients with subthreshold mania symptoms; MDD-, major depressive disorder patients with no subthreshold mania symptoms or risk factors for BD; AUC, area under the receiver operating characteristics (ROC) curve; Balanced Acc., balanced accuracy; PPV, positive predictive value; NPV, negative predictive value

(BDD & MDD+) versus MDD-. Discrimination maps showed that the temporal and parietal brain regions contributed most to these classifiers as shown in Fig. 2. In addition, the left amygdala region contributed to all these classifications of the MDD- group for all pairs.

Change in classification of MDD group over follow-up: Three MDD subjects (one MDD+ and two MDD-) converted to BD diagnosis over the course of follow-up (Koirala et al., 2019). We examined the original diagnostic label and the predicted diagnosis by classifier explained in the section multivariate pattern classification. One MDD- subject who converted was classified as BDD group in the classifier between BDD and MDD- for the happy face inhibition vs. non-emotional inhibition contrast. Another MDD- subject who converted was classified as the MDD+ group in both the classifiers for happy face inhibition vs. non-emotional inhibition contrast and emotional face inhibition vs. non-emotional inhibition contrast between MDD+ and MDD-. The third subject who converted was an MDD+ subject who was also classified as an MDD+ subject.

Discussion

In this first study comparing medication-free young depressed subjects with and without subthreshold bipolar symptoms and further comparison with young bipolar subjects and healthy controls, we found that MDD- group exhibited a different activation pattern compared to BDD and MDD+ groups. The differences were significant and extensive when emotional response conditions (happy, fearful, and combined) were contrasted with non-emotional response condition (gender). Post-hoc analysis of the main effect of group revealed that in the MDD- group compared to the BDD and MDD+ groups, exhibited a lower activation in areas of the parietal, temporal and frontal cortices during emotional



Region of Interest	Weight
Left Amygdala	-0.075622
Right Cerebellum (IV-V)	0.049623
Right Middle Occipital gyrus	-0.048331
Right Medial Temporal Pole	0.048234
Right Angular Gyrus	0.037013

Region of Interest	Weight
Right Middle Cingulate Cortex	-0.079947
Right Angular Gyrus	0.068734
Right Middle Temporal Gyrus	0.049236
Right Inferior Temporal Gyrus	-0.04102
Right Middle frontal Gyrus	0.024243

Region of Interest	Weight
Right Middle Cingulate Cortex	-0.049870
Left Amygdala	-0.036476
Right Medial Temporal Pole	0.034463
Right Precuneus	0.034299
Right Cerebellum (VI)	-0.034090

Region of Interest	Weight
Left Rectal Gyrus	-0.046170
Right Temporal Pole	0.041134
Right Superior Frontal Gyrus	0.040315
Left Amygdala	-0.035157
Right Precuneus	0.030348

Region of Interest	Weight
Right Cerebellum (IV-V)	0.070298
Left Temporal Pole	0.051128
Left Parahippocampal Gyrus	0.040474
Left Amygdala / Hippocampus	-0.035001
Left Inferior Frontal Gyrus	0.033538

Fig. 2. Receiver Operating Characteristics (ROC) Curves with Area Under Curve (AUC) and color weighted discrimination maps of classification for emotional face inhibition vs. non-emotional face inhibition generated by classifier using beta images masked by average effect of condition mask and top five most weighted brain regions for classification. Bonferroni correction was used to account for the multiple pairwise group comparisons. Results were considered significant if the corrected p-value for the balanced accuracy of classification < .05.
(A) BDD vs. MDD- classification. Positive weights (red) represent the voxels contributing to classification as a BDD subject, while negative weights (blue) represent the voxels contributing to classification as an MDD- subject. (B) BDD vs. MDD+ classification. Positive class: BDD subject, negative class: MDD+ subject. (C) MDD+ vs. MDD- classification. Positive class: MDD+ subject, negative class: MDD- subject. (D) (BDD & MDD+) vs. MDD- classification. Positive class: (BDD & MDD+) subject, negative class: MDD- subject. (E) BDD vs. (MDD+ & MDD-) classification. Positive class: BDD subject, negative class: (MDD+ & MDD-) subject.

(happy, fear, and combined emotional) response-inhibition vs. non-emotional response-inhibition (Table 1). Importantly, the BDD and MDD+ groups were not significantly different from each other.

The cuneus and adjacent areas of the parietal cortex and the temporal cortex are involved in the cognitive appraisal of emotion (Lettieri et al., 2019). Furthermore, the inferior frontal gyrus and cerebellum where lower activations were also seen are associated with motor response-inhibition to emotional stimuli (Simmonds, Pekar, & Mostofsky, 2008; Wessa et al., 2007). Response-inhibition to emotional stimuli in MDD- subjects was associated with the lowering of activation in the parieto-temporal-frontal emotional appraisal areas. The reason for this widespread deactivation in depression can only be speculated upon. For Go/No-go response-inhibition and other tasks in which attentional effort is required, studies have reported that either increased or decreased activations of brain areas involved with a successful performance on the task may be seen (Hester et al., 2004; Lawrence, Ross, Hoffmann, Garavan, & Stein, 2003). The decrease in activation may be necessary to suppress areas of the brain, which are overactive due to the monitoring of negative internal mood states. The BDD and MDD+ groups were also depressed and would have been expected to show a similar pattern to MDD-subjects. The higher activation in BDD and MDD+ seen in the parieto-temporal-frontal cortex suggests possible increased emotional reactivity, particularly to happy emotional stimuli in BDD and MDD+ groups, as has also been reported in other studies of fMRI activation to emotional stimuli in BD (J. R. C. d. Almeida et al., 2009; Blumberg et al., 2005; Fournier et al., 2013; Marchand et al., 2011).

Importantly, in accordance with the aim and hypothesis of this study, we were able to identify a different pattern of activation in BDD and MDD+ groups compared to the MDD- group. Therefore, the pattern could potentially be used to accurately classify BDD and MDD- and MDD+ and MDD- subjects. Indeed, machine learning classification of the different patient groups was possible using this difference of activation during emotional inhibition.

Activation related to happy face and combined emotion face response-inhibition vs. non-emotional face inhibition showed efficient classification (AUC: .73 and .70) between BDD and MDD- subjects. Combined emotional face response-inhibition vs. non-emotional face inhibition also showed good classification properties between MDD+ and MDD- groups (AUC: .70). This classification performance suggests that fMRI activation in response to emotional stimuli could potentially serve as a discriminator for identifying MDD+ subjects. Generally, an AUC of 0.7 to 0.8 is considered a suitable classification performance, 0.8 to 0.9 is considered excellent classification performance, and more than 0.9 is considered outstanding classification performance (Mandrekar, 2010). This study's classification results will fall under suitable classification and future studies with more discriminating task designs and a larger group of subjects will need to be conducted.

In combined emotional vs. non-emotional inhibition contrast, notably, the classification contribution of the MDD- group was weighted for the amygdala activation pattern (Fig. 1). Notably, in the univariate analysis, no differences in amygdala activation pattern were found. This finding underscores the importance of multivariate analysis, which can reveal findings not seen with univariate analysis. Abnormal amygdala activation has been reported in MDD (Anand et al., 2005) and studies have reported differing activation patterns in MDD and BD (J. R. C. Almeida, Versace, Hassel, Kupfer, & Phillips, 2010; Korgaonkar et al., 2019).

Notably, the post-hoc results of the main effects of group were not able to differentiate between BDD and MDD+ groups but there were several instances of significant differences between MDD+ and MDD- groups (Table 1). These findings indicated that the MDD+ group might be more similar to the BDD group than the MDD- group. Further support was found due to stronger discrimination between the combined BDD and MDD+ group compared to the MDD- group (AUC: .70 and .71 for happy and combined emotion vs. non-emotional response-inhibition) compared to the BDD group and the combined MDD+ and MDD- groups

(traditionally termed unipolar depression) (all AUCs < .70). These findings underscore the importance of studying MDD+ and MDD- groups separately when comparing bipolar and unipolar depression (de Almeida & Phillips, 2013).

There are several limitations of the current study. While the medication-free status of the participants in this study is a rare strength, it also poses possible selection bias in that the participant population may over-represent subjects with more mild conditions, and under-represent those where the severity of the disorder is such that medication-free study inclusion would be unethical. Because medication-free psychiatric neuroimaging is greatly understudied, these data are valuable but may lack full generalizability to medicated patient populations. These results must be compared with future studies in medicated and unmedicated participants to be appropriately generalized. For the same reason, we were unable to separate the BDD group into bipolar I and II subtypes for similarities and differences with the MDD+ group. However, we did not find any significant differences between BDI and BDII when the BD group was split in the LME analysis. However, a larger number of subjects in each group need to be studied. In future studies with a larger sample of bipolar subtypes, such an investigation will need to be undertaken. Furthermore, the emotional inhibition tasks in fMRI experiments can cause a cross-emotion effect. For example, participants may maintain the neuroimaging effects of happy facial stimuli when starting fear facial stimuli. This is a limitation of fMRI tasks. We used Neutral Go as the control stimulus to avoid responses to emotional stimulus, but cross-emotion effect cannot be fully eliminated. In future studies, fMRI paradigms which control for cross-emotion effects will need to be implemented.

For the classification analyses, ideally, a totally separate test-dataset from the training data set should be used. In future studies, the classifier developed in this study can be tested using a new sample. Finally, the greatest utility of classifiers such as that used in this study would be to predict the development of BD over time in MDD subjects. In this study, among three MDD subjects who converted to BD, two of the converted MDD- subjects were classified as BDD and the other converted MDD+ subject was classified as MDD+. That it was two MDD- subjects who converted indicates that some young MDD subjects not yet exhibiting subthreshold mania symptoms can also develop BD over time and that these subjects may be identified using the emotional Go-NoGo classifier. Though these results provide some face validity to the classifiers much more work needs to be done. Future studies with a larger number of subjects, a more extended follow-up period, and larger conversion rates will need to be conducted to validate the classifier.

Conclusion: The findings of the study indicate that the MDD- group was different from both BDD as well as MDD+ young adult groups on the regional activation pattern of emotional vs. non-emotional response inhibition tasks. MDD+ and BDD groups were more similar to each other. Future studies need to be conducted to further refine this classifier to predict the development of bipolar disorder in young adults presenting with depression.

Funding and disclosures

The NIMH funded this project to AA (R01MH093420). Dr. Anand, Dr. Altinay, Dr. Koirala, Sidra Speaker, Jungwon Cha, Jeffrey Spielberg, Bo Hu, Harish Karne, Amy Kuceyeski, and Elvisha Dhamala each reported no financial interests or potential conflicts of interest.

Contributors

Amit Anand conceived of the presented idea. Jungwon Cha, Jeffrey M. Spielberg, Harish Karne and Amit Anand conducted the experiment. Jungwon Cha, Sidra Speaker, Bo Hu, Murat Altinay, Parashar Koirala, Harish Karne, Jeffrey M. Spielberg, Amy Kuceyeski, Elvisha Dhamala and Amit Anand analyzed the data. All authors discussed the results and contributed to the final manuscript.

Disclosure

This article has been submitted to a preprint server BioRxiv on April 25, 2020, available online: https://doi.org/10.1101/2020.04.25.060111.

Acknowledgement

This study was funded by NIMH R01MH093420.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.jad.2020.12.037.

References

Almeida, J.R.C., Versace, A., Hassel, S., Kupfer, D.J., Phillips, M.L., 2010. Elevated Amygdala Activity to Sad Facial Expressions: A State Marker of Bipolar but Not Unipolar Depression. Biological Psychiatry 67 (5), 414–421. https://doi.org/10.1016/j.biopsych.2009.09.027.

Almeida, J.R.C.d., Versace, A., Mechelli, A., Hassel, S., Quevedo, K., Kupfer, D.J., Phillips, M.L., 2009. Abnormal Amygdala-Prefrontal Effective Connectivity to Happy Faces Differentiates Bipolar from Major Depression. Biological psychiatry 66 (5), 451–459. https://doi.org/10.1016/j.biopsych.2009.03.024.

Etkin, Amit, M.D., Ph.D., 2019. A Reckoning and Research Agenda for Neuroimaging in Psychiatry. American Journal of Psychiatry 176 (7), 507–511. https://doi.org/10.1176/appi.ajp.2019.19050521.

Anand, A., Li, Y., Wang, Y., Wu, J., Gao, S., Bukhari, L., Lowe, M., 2005. Activity and connectivity of brain mood regulating circuit in depression: A functional magnetic resonance study. Biological Psychiatry 57 (10), 1079–1088. https://doi.org/10.1016/j.biopsych.2005.02.021.

Angst, J., Gamma, A., Benazzi, F., Ajdacic, V., Eich, D., Rössler, W., 2003. Toward a re-definition of subthreshold bipolarity: epidemiology and proposed criteria for bipolar-II, minor bipolar disorders and hypomania. Journal of Affective Disorders 73 (1), 133–146. https://doi.org/10.1016/S0165-0327(02)00322-1.

Blumberg, H.P., Donegan, N.H., Sanislow, C.A., Collins, S., Lacadie, C., Skudlarski, P., Krystal, J.H., 2005. Preliminary evidence for medication effects on functional abnormalities in the amygdala and anterior cingulate in bipolar disorder. Psychopharmacology (Berl) 183 (3), 308–313. https://doi.org/10.1007/s00213-005-0156-7.

Brodersen, K.H., Haiss, F., Ong, C.S., Jung, F., Tittgemeyer, M., Buhmann, J.M., Stephan, K.E., 2011. Model-based feature construction for multivariate decoding. Neuroimage 56 (2), 601–615. https://doi.org/10.1016/j.neuroimage.2010.04.036.

Cardoso de Almeida, J.R., Phillips, M.L., 2013. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biological psychiatry 73 (2), 111–118. https://doi.org/10.1016/j.biopsych.2012.06.010.

Chen, G., Saad, Z.S., Britton, J.C., Pine, D.S., Cox, R.W., 2013. Linear mixed-effects modeling approach to FMRI group analysis. Neuroimage 73, 176–190. https://doi.org/10.1016/j.neuroimage.2013.01.047.

Chen, G., Taylor, P.A., Cox, R.W., 2017. Is the statistic value all we should care about in neuroimaging. Neuroimage 147, 952–959. https://doi.org/10.1016/j.neuroimage.2016.09.066.

Coryell, W., Endicott, J., Maser, J.D., Keller, M.B., Leon, A.C., Akiskal, H.S., 1995. Long-term stability of polarity distinctions in the affective disorders. The American journal of psychiatry 152 (3), 385–390. https://doi.org/10.1176/ajp.152.3.385.

Cox, R.W., Chen, G., Glen, D.R., Reynolds, R.C., Taylor, P.A., 2017. FMRI Clustering in AFNI: False-Positive Rates Redux. Brain Connect 7 (3), 152–171. https://doi.org/10.1089/brain.2016.0475.

de Almeida, J.R.C., Phillips, M.L., 2013. Distinguishing between Unipolar Depression and Bipolar Depression: Current and Future Clinical and Neuroimaging Perspectives. Biological Psychiatry 73 (2), 111–118. https://doi.org/10.1016/j.biopsych.2012.06.010.

Fiedorowicz, J.G., Endicott, J., Leon, A.C., Solomon, D.A., Keller, M.B., Coryell, W.H., 2011. Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder. The American journal of psychiatry 168 (1), 40–48. https://doi.org/10.1176/appi.ajp.2010.10030328.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33 (5), 636–647. https://doi.org/10.1002/mrm.1910330508.

Fournier, J.C., Keener, M.T., Mullin, B.C., Hafeman, D.M., Labarbara, E.J., Stiffler, R.S., Phillips, M.L., 2013. Heterogeneity of amygdala response in major depressive disorder: the impact of lifetime subthreshold mania. Psychol Med 43 (2), 293–302. https://doi.org/10.1017/s0033291712000918.

Frangou, S., Dima, D., Jogia, J., 2017. Towards person-centered neuroimaging markers for resilience and vulnerability in Bipolar Disorder. Neuroimage 145 (Pt B), 230–237. https://doi.org/10.1016/j.neuroimage.2016.08.066.

Golland, P., Fischl, B., 2003. Permutation tests for classification: towards statistical significance in image-based studies. Inf Process Med Imaging 18, 330–341.

Han, K.M., De Berardis, D., Fornaro, M., Kim, Y.K., 2019. Differentiating between bipolar and unipolar depression in functional and structural MRI studies. Prog Neuropsychopharmacol Biol Psychiatry 91, 20–27. https://doi.org/10.1016/j.pnpbp.2018.03.022.

Hester, R.L., Murphy, K., Foxe, J.J., Foxe, D.M., Javitt, D.C., Garavan, H., 2004. Predicting success: patterns of cortical activation and deactivation prior to response inhibition. J Cogn Neurosci 16 (5), 776–785. https://doi.org/10.1162/089892904970726.

Hummer, T.A., Hulvershorn, L.A., Karne, H.S., Gunn, A.D., Wang, Y., Anand, A., 2013. Emotional response inhibition in bipolar disorder: a functional magnetic resonance imaging study of trait- and state-related abnormalities. Biological psychiatry 73 (2), 136–143. https://doi.org/10.1016/j.biopsych.2012.06.036.

Kling, L.R., Bessette, K.L., DelDonno, S.R., Ryan, K.A., Drevets, W.C., McInnis, M.G., Langenecker, S.A., 2018. Cluster analysis with MOODS-SR illustrates a potential bipolar disorder risk phenotype in young adults with remitted major depressive disorder. Bipolar Disord 20 (8), 697–707. https://doi.org/10.1111/bdi.12693.

Koirala, P., Hu, B., Altinay, M., Li, M., DiVita, A.L., Bryant, K.A., Anand, A., 2019. Sub-threshold bipolar disorder in medication-free young subjects with major depression: Clinical characteristics and antidepressant treatment response. J Psychiatr Res 110, 1–8. https://doi.org/10.1016/j.jpsychires.2018.12.006.

Korgaonkar, M.S., Erlinger, M., Breukelaar, I.A., Boyce, P., Hazell, P., Antees, C., Malhi, G.S., 2019. Amygdala Activation and Connectivity to Emotional Processing Distinguishes Asymptomatic Patients With Bipolar Disorders and Unipolar Depression. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 4 (4), 361–370. https://doi.org/10.1016/j.bpsc.2018.08.012.

Lawrence, N.S., Ross, T.J., Hoffmann, R., Garavan, H., Stein, E.A., 2003. Multiple neuronal networks mediate sustained attention. J Cogn Neurosci 15 (7), 1028–1038. https://doi.org/10.1162/089892903770007416.

Lettieri, G., Handjaras, G., Ricciardi, E., Leo, A., Papale, P., Betta, M., Cecchetti, L., 2019. Emotionotopy in the human right temporo-parietal cortex. Nature Communications 10 (1), 5568. https://doi.org/10.1038/s41467-019-13599-z.

Mandrekar, J.N., 2010. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. Journal of Thoracic Oncology 5 (9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d.

Marchand, W.R., Lee, J.N., Garn, C., Thatcher, J., Gale, P., Kreitschitz, S., Wood, N., 2011. Aberrant emotional processing in posterior cortical midline structures in bipolar II depression. Prog Neuropsychopharmacol Biol Psychiatry 35 (7), 1729–1737. https://doi.org/10.1016/j.pnpbp.2011.05.017.

Merikangas, K.R., Jin, R., He, J.P., Kessler, R.C., Lee, S., Sampson, N.A., Zarkov, Z., 2011. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry 68 (3), 241–251. https://doi.org/10.1001/archgenpsychiatry.2011.12.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15 (1), 1–25.

Noirhomme, Q., Lesenfants, D., Gomez, F., Soddu, A., Schrouff, J., Garraux, G., Laureys, S., 2014. Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. Neuroimage Clin 4, 687–694. https://doi.org/10.1016/j.nicl.2014.04.004.

Nurnberger, J.I., Jr., McInnis, M., Reich, W., Kastelic, E., Wilcox, H.C., Glowinski, A., Monahan, P.O., 2011. A high-risk study of bipolar disorder. Childhood clinical phenotypes as precursors of major mood disorders. Arch Gen Psychiatry 68 (10), 1012–1020. https://doi.org/10.1001/archgenpsychiatry.2011.126.

Phillips, M.L., Swartz, H.A., 2014. A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. Am J Psychiatry 171 (8), 829–843. https://doi.org/10.1176/appi.ajp.2014.13081008.

Piguet, C., Fodoulian, L., Aubry, J.M., Vuilleumier, P., Houenou, J., 2015. Bipolar disorder: Functional neuroimaging markers in relatives. Neurosci Biobehav Rev 57, 284–296. https://doi.org/10.1016/j.neubiorev.2015.08.015.

Rasmussen, C. E. (2003). *Gaussian processes in machine learning*. Paper presented at the Summer School on Machine Learning.

Redlich, R., Almeida, J.J., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., Dannlowski, U., 2014. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. JAMA Psychiatry 71 (11), 1222–1230. https://doi.org/10.1001/jamapsychiatry.2014.1100.

Redlich, R., Dohm, K., Grotegerd, D., Opel, N., Zwitserlood, P., Heindel, W., Dannlowski, U., 2015. Reward Processing in Unipolar and Bipolar Depression: A Functional MRI Study. Neuropsychopharmacology 40 (11), 2623–2631. https://doi.org/10.1038/npp.2015.110.

Simmonds, D.J., Pekar, J.J., Mostofsky, S.H., 2008. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46 (1), 224–232. https://doi.org/10.1016/j.neuropsychologia.2007.07.015.

Strakowski, S.M., Adler, C.M., Almeida, J., Altshuler, L.L., Blumberg, H.P., Chang, K.D., Townsend, J.D., 2012. The functional neuroanatomy of bipolar disorder: a consensus model. Bipolar Disord 14 (4), 313–325. https://doi.org/10.1111/j.1399-5618.2012.01022.x.

Townsend, J.D., Bookheimer, S.Y., Foland-Ross, L.C., Moody, T.D., Eisenberger, N.I., Fischer, J.S., Altshuler, L.L., 2012. Deficits in inferior frontal cortex activation in euthymic bipolar disorder patients during a response inhibition task. Bipolar Disord 14 (4), 442–450. https://doi.org/10.1111/j.1399-5618.2012.01020.x.

Vierck, E., Porter, R.J., Joyce, P.R., 2015. Facial recognition deficits as a potential endophenotype in bipolar disorder. Psychiatry Res 230 (1), 102–107. https://doi.org/10.1016/j.psychres.2015.08.033.

Wessa, M., Houenou, J., Paillere-Martinot, M.L., Berthoz, S., Artiges, E., Leboyer, M., Martinot, J.L., 2007. Fronto-striatal overactivation in euthymic bipolar patients during an emotional go/nogo task. Am J Psychiatry 164 (4), 638–646. https://doi.org/10.1176/ajp.2007.164.4.638.

Zimmermann, P., Brückl, T., Nocon, A., Pfister, H., Lieb, R., Wittchen, H.-U., Angst, J., 2009. Heterogeneity of DSM-IV major depressive disorder as a consequence of subthreshold bipolarity. Archives of general psychiatry 66 (12), 1341–1352. https://doi.org/10.1001/archgenpsychiatry.2009.158.