



# Cortical thickness predicts the first onset of major depression in adolescence



Lara C. Foland-Ross<sup>a,\*,1</sup>, Matthew D. Sacchet<sup>a,b,1</sup>, Gautam Prasad<sup>c</sup>, Brooke Gilbert<sup>a</sup>, Paul M. Thompson<sup>c</sup>, Ian H. Gotlib<sup>a,b</sup>

<sup>a</sup> Psychology Department, Stanford University, CA, USA

<sup>b</sup> Neurosciences Program, Stanford University, CA, USA

<sup>c</sup> Keck School of Medicine at the University of Southern California, CA, USA

## ARTICLE INFO

### Article history:

Received 26 May 2015

Received in revised form 18 July 2015

Accepted 20 July 2015

Available online 24 August 2015

### Keywords:

Depression

Cortical thickness

Emotion regulation

Machine learning

## ABSTRACT

Given the increasing prevalence of Major Depressive Disorder and recent advances in preventative treatments for this disorder, an important challenge in pediatric neuroimaging is the early identification of individuals at risk for depression. We examined whether machine learning can be used to predict the onset of depression at the individual level. Thirty-three never-disordered adolescents (10–15 years old) underwent structural MRI. Participants were followed for 5 years to monitor the emergence of clinically significant depressive symptoms. We used support vector machines (SVMs) to test whether baseline cortical thickness could reliably distinguish adolescents who develop depression from adolescents who remained free of any Axis I disorder. Accuracies from subsampled cross-validated classification were used to assess classifier performance. Baseline cortical thickness correctly predicted the future onset of depression with an overall accuracy of 70% (69% sensitivity, 70% specificity;  $p = 0.021$ ). Examination of SVM feature weights indicated that the right medial orbitofrontal, right precentral, left anterior cingulate, and bilateral insular cortex contributed most strongly to this classification. These findings indicate that cortical gray matter structure can predict the subsequent onset of depression. An important direction for future research is to elucidate mechanisms by which these anomalies in gray matter structure increase risk for developing this disorder.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Major Depressive Disorder (MDD) is the most common psychiatric illness in the United States (Kessler et al., 2003) and accounts for almost half of all disability-adjusted life years worldwide (Whiteford et al., 2013). Although the onset of major depression can occur at any age, most adults living with the disorder experienced their first episode during their teenage years (Lewinsohn et al., 1998). Late adolescence is a period of considerable vulnerability for depression: one-quarter of all youth will experience an episode of MDD by the end of their teenage years, a figure that is increasing over time (Kessler et al., 2001). Despite over two decades of research examining the neural and molecular bases of depres-

sion, fewer than one-quarter of depressed adolescents respond to an initial course of antidepressant medication. Half of those who do respond to treatment continue to have disturbances in sleep and mood, fatigue, and concentration difficulties (Tao et al., 2010).

Given the enormous personal and societal costs and the high prevalence of MDD, and the difficulty in treating depression once it has developed, early detection and prevention is crucial. Prevention programs that are implemented before the first onset of disorder should significantly reduce the societal burden of adolescent and adult depression. Such programs, however, require that we accurately identify children at greater risk for the disorder (Gladstone et al., 2011), a goal that continues to be elusive as the etiology of major depression is poorly understood. Therefore, an important challenge in pediatric neuroimaging is to identify brain measures that are useful for predicting the onset of MDD. In this study, we used machine learning methods to test the hypothesis that structural patterns in cortical gray matter can distinguish adolescents who go on to develop clinically significant depression within five years of their initial assessment from adolescents who remain free of any Axis I disorder. As in previous studies (Grotegerd et al., 2013;

\* Corresponding author at: Stanford Mood and Anxiety Disorders Laboratory, Department of Psychology, Stanford University, 450 Serra Mall, Stanford, Jordan Hall, Building 420, CA 94305, USA.

E-mail address: [lfolandross@stanford.edu](mailto:lfolandross@stanford.edu) (L.C. Foland-Ross).

<sup>1</sup> These authors contributed equally to this study.

**Table 1**  
Demographic variables.

Demographic variable	Converted (N = 18)	Non-converted (N = 15)	p
Age at baseline (SD), years	13.0 (1.8)	13.7 (1.2)	0.25 <sup>a</sup>
Baseline CDI-S score (SD)	0.83 (1.5)	0.73 (0.8)	0.81 <sup>a</sup>
Baseline MASC score (SD)	39.3 (9.8)	40.5 (13.8)	0.77 <sup>a</sup>
Baseline WISC-III vocabulary score (SD)	46.8 (5.1)	50.4 (6.7)	0.11 <sup>a</sup>
Baseline SES	3.8 (1.5)	4.2 (0.7)	0.11 <sup>b</sup>
Maternal history of MDD (%)	9 (50.0)	11 (73.3)	0.28 <sup>b</sup>
Number of traumatic events experienced before scanning	1.2 (0.9)	0.9 (1.1)	0.36 <sup>a</sup>
Number of traumatic events experienced between baseline and follow-up assessment	1.6 (1.0)	1.9 (1.6)	0.84 <sup>a</sup>

Values indicate the mean  $\pm$  SD unless otherwise noted.

Abbreviations: CDI-S, Child Depression Inventory–Short Form; MASC, Multidimensional Anxiety Scale for Children; WISC-III, Wechsler Intelligence Scale for Children–III; SES, socioeconomic status; MDD, Major Depressive Disorder.

<sup>a</sup> Statistic computed using two-sample *t*-tests.

<sup>b</sup> Statistic computed using  $\chi^2$  test, Trauma was assessed using Post Traumatic Stress Disorder (PTSD) section of the K-SADS-PL. SES was coded using annual household income (US dollars) as follows: 0 = less than 10,000, 1 = 10,000–25,000, 2 = 25,000–50,000, 3 = 50,000–75,000, 4 = 75,000–100,000, 5 = more than 100,000.

Redlich et al., 2014), our SVM analysis was conducted using a mask that included areas of the cortex that have been widely implicated in neuroanatomical models of depression and of emotion regulation (e.g., ventral and superior frontal cortex, anterior cingulate cortex, insula).

## 2. Methods

### 2.1. Participants

The study was approved by Stanford University's institutional review board. All adolescent participants provided written assent, and their mothers provided written informed consent. Participants in the current study sample included 33 adolescent girls, ages 10–15 years. We restricted our sampling to young adolescent girls to avoid the confound of sex differences on brain structure and depression vulnerability (Nolen-Hoeksema & Hilt, 2010) and because of previous findings indicating that the average age for first onset of adolescent depression is 15 years (Lewinsohn et al., 1994). Participants were recruited through advertisements posted in numerous locations (e.g., internet bulletin boards, university kiosks, supermarkets, etc.) describing a research study on depression in mothers. The mothers' responses to a telephone interview provided initial selection information. This phone screen established that both mothers and daughters were fluent in English and that daughters were between 10 and 15 years of age. Participants were excluded if they had experienced severe head trauma, learning disabilities, current or past depression or other Axis I disorders, current or past substance abuse, or if they were taking psychotropic medications. The telephone interview was also used to identify mothers who were likely either to have no psychiatric history or to meet criteria for recurrent depression during their daughter's lifetime. Those mother–daughter pairs who were considered likely to be eligible for participation were invited to come to the laboratory.

Baseline diagnostic status of participants was assessed by administering the structured Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (K-SADS-PL; Kaufman et al., 1997) to both the girls and their mothers (regarding the girls). In addition, the Structured Clinical Interview for the DSM-IV was administered to participants' mothers to assess history of maternal depression (for more information about the use of these measures in this study, see Joormann et al., 2007; Gotlib et al., 2010; Chen et al., 2012). To monitor the course of depressive symptoms and the possible onset of depression, girls and their mothers returned to the laboratory for diagnostic reassessment every 18 months for up to 5 years, or until the onset of depression. Importantly, although investigators have estimated that between 12% (Merikangas et al., 2010)

and 35% (Weissman et al., 1997) of adolescents will experience an initial onset of depression by the end of their teenage years, just over 50% of the participants in the current study had developed a depressive disorder by follow-up (see Section 3). This high incidence rate was due to our sampling approach: many of the adolescents were at familial risk for depression by virtue of having a mother with repeated episodes of depression during her lifetime (Table 1). More importantly, however, to qualify for inclusion in the converted group, participants needed to be followed only until the onset of a depressive episode; this took considerably less time from the baseline scan date than the time needed to qualify as for inclusion in the non-converted group (i.e., had not developed a depressive episode in the five years following the baseline scan date).

#### 2.1.1. Magnetic resonance imaging data acquisition

Within one week of the baseline interview assessments, participants were scanned on a 1.5T GE scanner (GE Healthcare Systems, Milwaukee, Wisconsin). Anatomic images were obtained using a T1-weighted spoiled gradient-recalled (SPGR) echo sequence using the following parameters: repetition time (TR) = 8.924 ms, echo time (TE) = 1.792 ms, flip angle = 15°, in-plane resolution = 0.859  $\times$  0.859 mm<sup>2</sup>, and slice thickness = 1.5 mm.

#### 2.1.2. Magnetic resonance imaging data preprocessing

Scans were processed using FreeSurfer to produce measures of cortical gray matter thickness (Fischl & Dale, 2000; version 5.0, <http://surfer.nmr.mgh.harvard.edu>; Huang and Du, 2005; Salat et al., 2004). Processing streams included the removal of nonbrain tissue, intensity normalization, the segmentation of gray/white matter, and the alignment of each image volume to a standardized space. To estimate cortical gray matter thickness, a deformable surface algorithm was applied to segmented images to extract the pial and gray/white cortical surfaces (Dale et al., 1999). To ensure the accuracy of gray/white matter segmentation, exclusion of scalp and other non-brain tissue, and the inclusion of brain tissue, cortical surfaces were visually inspected by one rater (B.G.) who was blind to group membership. Manual corrections were performed by this rater where appropriate, in accordance with previously established procedures (Black et al., 2012; Yang et al., 2009). After spatially normalizing the data to an average template space, local cortical thickness was measured by estimating the shortest distance between spatially equivalent surface points on the pial surface and the gray–white matter boundary and vice versa and averaging these 2 values (Fischl and Dale, 2000). The cortical surface was then parcellated into 34 regions per hemisphere, as in Desikan et al. (2006). We chose to use cortical gray matter thickness as our primary measure in the current study, given prior work by our group and

others documenting cortical thickness abnormalities in the context of depression risk (Foland-Ross et al., *in press*; Peterson et al., 2009; Pappmeyer et al., 2014) and given evidence that this measure is less strongly influenced by individual variations in surface area than is volume (Winkler et al., 2010). However, investigations that use other aspects of brain structure (e.g., cortical surface area, gyrification index, cortical and subcortical volume) to predict subsequent onset of depression would also be informative in the broader understanding of a neural risk for this disorder.

### 2.1.3. Neuroanatomical classification

SVM is a tool from the field of machine learning that classifies novel data with a classifier that has been trained on data labeled with true class labels (Braver et al., 1997). SVM has been successful when applied to psychiatric and neurological diagnosis, identifying prodromal states, and in predicting treatment prognosis (for a review, see Orrù et al., 2012). In the current study, class labels included two categorical groups: girls who developed a diagnosable episode of depression within five years ( $N=18$ ; i.e., “converted”), and girls who did not develop any Axis I disorder within five years ( $N=15$ ; i.e., “non-converted”). The feature set, or the data used for classification, was composed of regional thickness averages for a subset of brain areas previously implicated in neuroanatomical models of emotion and depression (Grotegerd et al., 2013; Redlich et al., 2014). The mask was created using the Desikan atlas according to the automated anatomical labeling definitions (Desikan et al., 2006). As with previous SVM studies of mood disorders (Grotegerd et al., 2013; Redlich et al., 2014). This large region of interest, which included nearly the entire frontal lobe and some parts of the temporal lobe was selected with the goal of including those neural regions most likely to substantially contribute to discriminative patterns in order to increase the performance of the classifiers. The complete list of subregions included in the mask is noted in Supplementary material. Converted and non-converted girls in the current study did not differ with respect to age or intracranial volume ( $ts < 1.17$ ,  $ps > 0.25$ ). Moreover, exploratory analyses, conducted at the vertex level using a general linear model in QDEC (part of Freesurfer; <http://surfer.nmr.mgh.harvard.edu/fswiki/Qdec>) and carried out using regional thickness averages and a general linear model in SPSS (<http://www-01.ibm.com/software/analytics/spss>) indicated no significant or trend-level associations among cortical thickness, age, and total intracranial volume. Thus, the thickness averages used in our SVM were not adjusted for these factors. While the absence of an association between thickness and age may be surprising, at least one report has also noted a lack of such a relation (e.g., Pakkenberg and Gundersen, 1997). It is also possible, however, that the small sample size limited our power to detect such an association.

SVMs, by definition, find the hyperplane, or high-dimensional surface, that best differentiates two classes in a training data feature set. This differentiation is achieved by maximizing the “margin” between data points from the two classes (Cortes and Vapnik, 1995). The selected model can then be used to classify novel, or “test,” data points. Because the converted and non-converted groups did not contain equal sample sizes, which can bias classification estimates (Japkowicz and Stephen, 2002; Hastie et al., 2005), we repeatedly randomly subsampled the larger (converted,  $N=18$ ) group, such that it was the same size as the smaller group (non-converted,  $N=15$ ). For each subsampled dataset, we used stratified 10-fold cross-validation (Kohavi, 1995) to estimate generalized accuracy, sensitivity, and specificity of classification for each (Hastie et al., 2005). More specifically, for each of 10 folds, disjoint subsets were created for training ( $N=27$ ) and testing ( $N=3$ ) the classifier such that no individual was ever in both sets. Such an approach has been shown to improve bias and variance estimates relative to other cross-validation schemes (Kohavi, 1995).

To arrive at a stable estimate of classifier performance with this combined cross-validation and subsampling approach, we conducted 10-fold cross-validation 50 times with randomly identified subsamples and report average performance metrics across these subsamples. As described and implemented in Fu et al. (2008) and Redlich et al. (2014), we used the binomial test to assess whether classification performed above chance. Because the calculation of performance across subsamples could yield a non-integer value for performance (i.e., average number of successful classifications of 30 tests across subsamples), and the binomial distribution is discrete and defined for integers, we conducted binomial testing using rounded average performance across subsamples. To examine the features (i.e., brain regions) that contributed most strongly to classification, feature weights were computed according to the method described by De Martino et al. (2008). Briefly, this approach quantifies feature weights based on the weight of a given feature in the SVM-defined model (i.e., hyperplane), with greater weights indicating greater contribution to the model. We computed feature weight ranks using all individuals in a given subsample and then averaged these ranks across subsamples. Details of the derivation of the SVM-defined optimal hyperplane are presented in Supplementary material.

## 3. Results

### 3.1. Participant characteristics

Demographic and clinical characteristics of the participants are presented in Table 1. Eighteen of the 33 girls in this study (54.5%) subsequently developed depression (i.e., converted) within 5 years of the baseline scan (average age of onset =  $16.2 \pm 1.8$  years). Converted girls did not differ in age at baseline from non-converted girls, in their baseline scores on the Children's Depression Inventory—Short Form (CDI-S; Kovacs, 1992), the Multidimensional Anxiety Scale for Children (MASC; March et al., 1997), or the vocabulary subscale of the Wechsler Intelligence Scale for Children—III (WISC-III; Weschler, 1991). There were also no group differences in baseline socioeconomic status, maternal history of depression, assessed using the Structured Clinical Interview for DSM-IV (American Psychiatric Association, 2000) or in the number of traumatic events experienced either before baseline scanning or between baseline and follow-up assessments, as measured with the Post Traumatic Stress Disorder (PTSD) section of the K-SADS-PL.

Of the 18 participants who converted, 11 (61.1%) endorsed 5 or more symptoms of depression on the K-SADS-PL and met diagnostic criteria for MDD. The remaining 7 girls endorsed 3 (16.7%) or 4 (22.2%) symptoms of depression and met criteria for a diagnosis of Depressive Disorder Not Otherwise Specified (DD-NOS). With respect to comorbidities, of the 18 girls who converted, 10 (55.6%) developed one or more anxiety disorders. No girls in the non-converted group developed an anxiety disorder, and no other psychiatric comorbidities were documented in either group.

### 3.2. Neuroanatomic classification

Initial testing of classifier performance revealed an average generalized accuracy of 69.7% (69.3% sensitivity, 70.0% specificity) across cross-validated subsampled datasets. Statistical testing using the binomial test showed that classification accuracy was significantly better than chance ( $p=0.021$ ). Feature weight ranks indicated that the gray matter thickness of the right medial orbitofrontal cortex contributed most strongly to classification (i.e., exhibited the highest feature weight rank). Exploratory post-hoc two-tailed  $t$ -tests indicated that girls who subsequently developed depression had significantly thinner gray matter in this region than

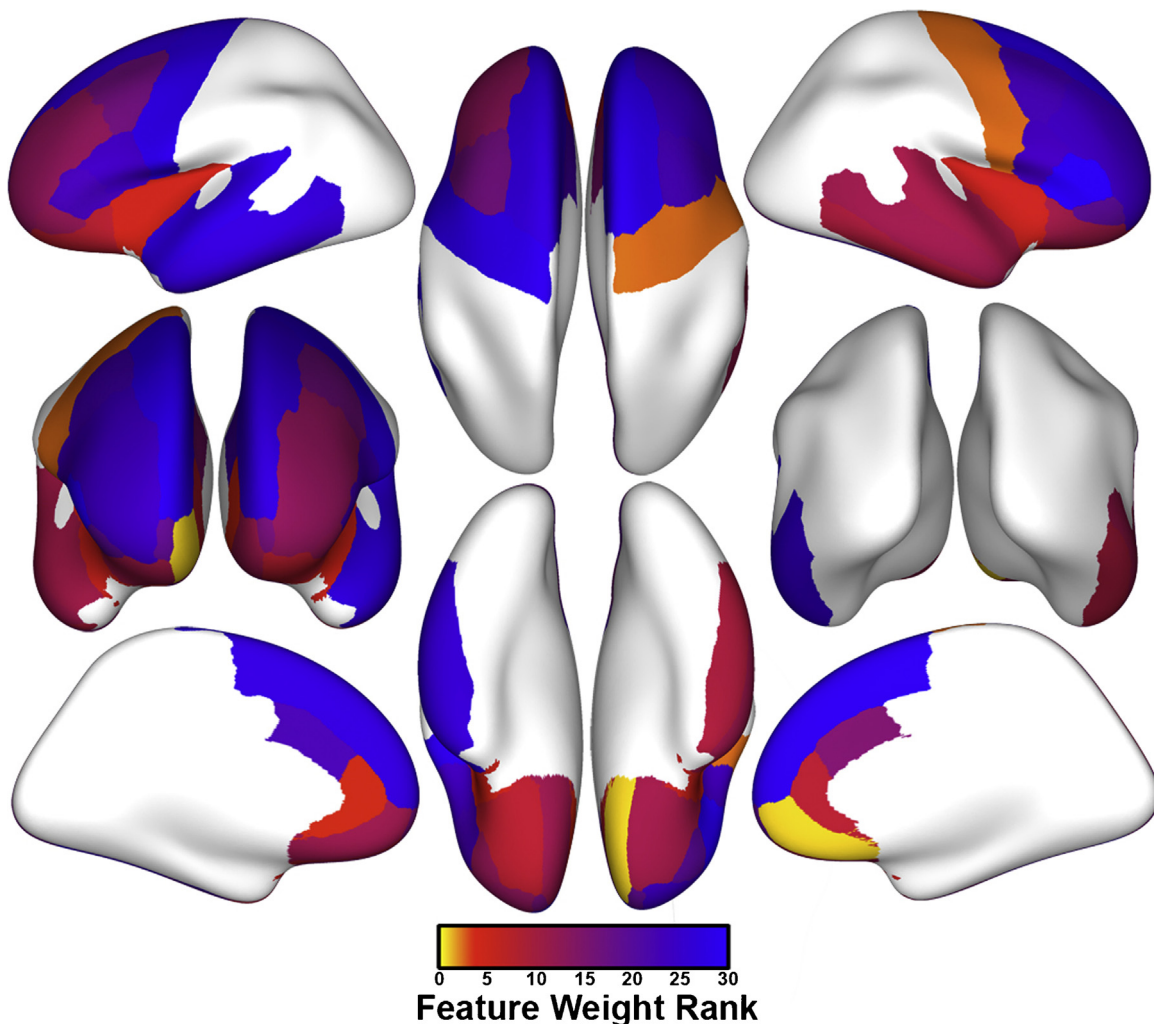
**Table 2**  
Ten features having the greatest contribution to classification accuracy in the SVM.

Overall Rank	Average Rank	Hemisphere	Average regional thickness (SD), mm		<i>p</i>	Region
			Non-converted	Converted		
1	1.00	Right	2.64 (0.18)	2.45 (0.15)	0.004	Medial orbitofrontal cortex
2	2.52	Right	2.58 (0.22)	2.64 (0.10)	0.351	Precentral gyrus
3	3.66	Left	2.84 (0.20)	2.95 (0.19)	0.137	Rostral anterior cingulate cortex
4	5.22	Left	2.95 (0.12)	3.09 (0.19)	0.015	Insula
5	6.10	Right	2.97 (0.19)	2.95 (0.22)	0.770	Insula
6	8.38	Left	2.53 (0.13)	2.57 (0.13)	0.312	Lateral orbital frontal cortex
7	8.52	Right	2.75 (0.25)	2.87 (0.20)	0.156	Rostral anterior cingulate cortex
8	9.80	Right	2.72 (0.13)	2.74 (0.18)	0.648	Inferior temporal gyrus
9	10.94	Right	2.64 (0.18)	2.66 (0.16)	0.717	Superior temporal gyrus
10	11.14	Right	2.60 (0.22)	2.48 (0.12)	0.051	Lateral orbital frontal cortex

Laterality, region name, and average feature weight rank of the highest ranked features are displayed. Regions having a lower overall rank contributed a greater amount of information to the classifier. Significance values indicate uncorrected *p* values from post-hoc comparisons of regional thickness averages between groups using two-tailed independent samples *t*-tests.

did girls who did not develop depression (average thickness: 2.45 versus 2.64 mm, respectively;  $t(31) = 3.15$ , uncorrected  $p = 0.004$ ). The 10 regions that contributed the most information to the classifier are presented in Table 2 (see Supplementary material for a complete listing of all 32 regions). As shown in this table, gray matter of the right precentral gyrus, the left rostral anterior cingulate cortex, and the left and right insula were among the top

five ranked features (Fig. 1, Table 2). Post-hoc two-tailed *t*-tests of these 5 areas indicated that gray matter of the left insula was significantly thicker in girls who converted than in girls who did not convert (average thickness: 3.09 mm versus 2.94 mm, respectively;  $t(31) = -2.56$ , uncorrected  $p = 0.015$ ); the two groups of girls did not differ significantly with respect to thickness in any other of the top 10 ranked regions ( $ts < 1.53$ ,  $ps > 0.14$ ).



**Fig. 1.** Average feature weight ranks across support vector machines (SVMs) from all individuals from each of 50 subsamples for all 32 cortical regions included in our anatomical mask. Features with the greatest contributions toward classifier accuracy (and thus the highest average feature weight ranks) are coded in yellow and orange.



#### 4. Discussion

Using prospective analyses and machine learning methods, we identified patterns of brain structure in adolescent girls that differentially predicted the subsequent onset of depression within five years. Specifically, the thickness of the orbitofrontal cortex, precentral gyrus, anterior cingulate, and insula were among the regions that provided the most information in distinguishing adolescents who went on to develop clinically significant depression from those who remained free of any Axis I disorder. Given the high costs of depression (Murray and Lobe, 1996), the difficulty in treating adolescent depression once it occurs (March et al., 2004), and the initial promise of early interventions (Gladstone et al., 2011), the set of biomarkers for the onset of depression identified here has the potential to make significant clinical implications. Indeed, distinguishing to determine which high-risk adolescents are at greatest risk for developing MDD and offering them prevention-focused intervention may significantly reduce the enormous negative personal and societal impact of this debilitating disorder.

The greatest contribution from any single brain region to classification came from the right medial orbitofrontal cortex. Therefore, structure in this area alone may serve as an important biomarker of the initial prodrome of depression. A role for the right medial orbitofrontal cortex in depression is well established. Alterations – primarily reductions – in medial orbitofrontal volume have been documented in several studies of chronically depressed adults (Bremner et al., 2002; Grieve et al., 2013; Koolschijn et al., 2009; Lacerda et al., 2004; Nakano et al., 2014). This direction of effects is consistent with post-hoc analyses of this region in the current study. Specifically, we found that girls who went on to develop depression at follow-up exhibited significant reductions in the thickness of this region at baseline relative to girls who remained well (uncorrected  $p = 0.004$ ). Alterations of the medial orbitofrontal cortex, therefore, likely precede the emergence of depression and may be involved in the genesis of the disorder. It is interesting to note that the single largest neuroimaging study of familial risk for depression found that this same area of the medial orbitofrontal cortex showed an opposite direction of effects. That is, in this prior study, adults at familial risk for depression showed abnormally thick right medial orbitofrontal gray matter (Peterson et al., 2009). How this prior pattern of findings is associated with our own is unclear but may be related to differential neural factors underlying risk versus resilience. Because the mean age of participants in this prior study of risk were significantly older than the age of peak incidence for depression (Weissman et al., 1997), cortical thickening of this region could have signified some form of resilience to the disorder. Certainly, future studies that scan high- and low-risk individuals both before and following the onset (or the non-onset) of depression are needed to test this interpretation.

The medial orbitofrontal cortex plays a central role in monitoring the rewarding value of stimuli and in anticipating upcoming emotional states or reactions (Amodio and Frith, 2006). Thus, structural anomalies of this region could contribute to the onset of depression by disrupting the neural processing of reward. Consistent with this formulation, anomalous structure and function of the orbitofrontal cortex is one of the most frequently reported findings in neuroimaging studies of anhedonia (Gorwood, 2008), which has been conceptualized both as a core symptom of depression (Pizzagalli, 2014) and as a risk factor for the disorder (Pizzagalli, 2014; Rawal et al., 2013). Indeed, low reward-seeking behavior in adolescents at familial risk for depression has been found to predict depressive symptoms and the onset of MDD one year later, even when controlling for baseline symptom levels (Rawal et al., 2013). We do not know whether the thinner gray matter in the right medial orbitofrontal region of the converted girls in this study is related to anomalous reward-seeking behavior; future neuroimag-

ing studies of developing adolescents that incorporate assessments of all of these constructs will be necessary to elucidate the behavioral correlates of the medial orbitofrontal abnormalities that we found in this study.

Exploratory analyses of the remaining nine features that contributed most strongly to classification accuracy in the SVM indicated that the converted and non-converted girls differed from each other in the thickness of only one region other than right medial orbitofrontal cortex: the left insula. This region, particularly the anterior segment, is central to the detection and processing of subjectively salient stimuli (Rolls, 1996) and to integrating this information for the initiation of appropriate control signals (Mennon & Uddin, 2010). The anterior insula is also involved in the experience of social emotions, such as guilt (Lamm & Singer, 2010) – a core symptom of depression in pediatric samples (Pagliaccio et al., 2014). Thus, it is possible that structural alterations in this area contribute to the onset of depression by influencing a disproportionate allocation of resources to the internal experience of emotion.

In contrast to our findings of thinner gray matter in the right medial orbitofrontal cortex, girls who subsequently developed a depressive disorder had greater thickness of the left insula than did girls who remained well. The direction of this effect is in contrast to a recent neuroimaging investigation of preschool children that reported smaller insula volumes to be associated with subsequent depression (Belden et al., 2015). Our finding that increased insula thickness predicted the onset of depression also contrasts with neuroimaging studies of depressed adults that report MDD-associated reductions in insula volume (Koolschijn et al., 2009). It is important to note, however, that several of the studies that have assessed brain structure in individuals scanned early in the course of their disorder have reported fronto-limbic *increases* in cortical gray matter associated with the initial onset of a depressive disorder (Qiu et al., 2014; Reynolds et al., 2014; Van Eijndhoven et al., 2013). Thus, although speculative given our prospective study design, when taken in context with prior neuroimaging studies of first-episode and chronically ill individuals, our findings in the insula suggest a pathological process whereby cortical gray matter of this region is higher prior to the onset of the disorder, but ultimately becomes thinner as individuals experience repeated episodes of depression. While longitudinal studies are clearly required to test this formulation, one recent longitudinal study of young adults at familial risk for bipolar disorder reported that the onset of mood disorder at follow-up was associated with anomalous increases over time in the thickness of neighboring gray matter of the inferior frontal gyrus (Pappmeyer et al., 2014). Adding support to a potential reversal in the direction of the association of insula anomalies and depression are studies showing that later in the course of the disorder, repeated episodes of depression are associated with a decline in gray matter in frontal lobe subregions. For example, Reynolds et al. (2014) found that depressed adolescents showed a stronger inverse relation between age and middle frontal thickness than did never-depressed controls. Similarly, Frodl et al. (2008) reported that adults who experienced one or more episodes of depression over a three-year time period exhibited greater gray matter reductions between baseline and follow-up scans than did adults who had remitted from depression. Investigations in which adolescents or young adults are followed longitudinally starting before the appearance of symptoms and continuing through the onset and course of depression, and in which other aspects of psychobiological functioning are assessed, will be critical in elucidating the temporal patterns of structural neural anomalies and their underlying causes.

A unique strength of the present study is the careful selection of the study sample. Although the onset of depression can occur at any age, the peak age of incidence ranges from 15 to 20 years (Weissman et al., 1997), a range immediately following the age at

which the current sample was scanned (10–15 years). It is also noteworthy that the converted and non-converted girls in this study did not differ with respect to either maternal history of depression or the number of traumatic events reported prior to scanning or between the baseline and follow-up assessments. A primary limitation of the current study, however, is the sample size. The number of participants included in the groups of converted and non-converted girls was relatively small, and our findings should be replicated with larger samples. Importantly however, the feature-to-individual ratio in the current study was relatively large (1.07), which, combined with our random subsampling approach, was helpful for modeling these groups and reducing bias in our classification estimates (Raudys and Jain, 1991; Redlich et al., 2014). We should also note that 10 of the 18 girls who developed a depressive disorder by follow-up also developed a comorbid anxiety disorder, and 7 girls endorsed symptoms of depression and met criteria for a diagnosis of Depressive Disorder Not Otherwise Specified (DD-NOS). Future studies with larger samples that include participants who subsequently develop non-comorbid depression and depression-NOS, non-comorbid anxiety, and comorbid depression and anxiety are needed to accurately assess the specificity of our classifier in predicting MDD, MDD-NOS, comorbid MDD/anxiety and non-comorbid anxiety. We can also not exclude the possibilities that girls in our non-converted sample developed depression after 5 years from their scan and that including such girls in our non-converted group affected classifier performance. Future studies that test whether classification accuracy is influenced by the length of time between scan and the subsequent first onset of depression would be helpful in addressing this issue. Although we did not conduct immediate repeated scans of the participants in the same session, which would be required to formally test the reliability of our measurement of cortical thickness, we did use established procedures to maximize the accuracy of cortical thickness measurements in Freesurfer; in addition, we ensured that raters were blind to group membership, and we used an identical scanner and similar acquisition parameters to those used by Dickerson et al. (2008), who found absolute measurements of cortical thickness to be highly reliable. Finally, because we focused in this study on structural neuroimaging predictors of subsequent depression, we cannot determine how variations in brain functioning might relate to a future onset of MDD. Given the growing literature regarding abnormalities in functional activations in adolescent depression (for review see Kerestes et al., 2014; Miller et al., 2015), it will be important in future research to examine whether patterns of neural activation predict onset of the first episode of depression.

In conclusion, using machine-learning methods we demonstrated that patterns of brain structure in adolescence can help in predicting the subsequent onset of depression. Our findings represent an important first step towards understanding the pathological neural processes that underlie the emergence of this disorder and highlight several important avenues for future research. First and foremost, given our relatively small sample size, it is important that future studies, using larger samples and alternate cross-validation designs, replicate the anomalies in medial orbitofrontal and insular gray matter thickness found here. It is also important that future studies attempt to clarify the pathophysiological processes at play in the development of disorder. Such studies could, for example, assess whether alterations in cortical thickness influence the onset of depression directly or indirectly by mediating the emergence of risk-related biases in cognitive functioning. It is imperative that we continue to address this and other questions concerning the prediction of MDD in pediatric populations using multivariate approaches. This effort will elucidate biomarkers that can better identify individuals at high risk for developing depression in order to reduce the economic, societal, and personal costs of this recurrent and debilitating disorder.

## Conflict of Interest

None declared.

## Acknowledgements

We thank all the participants and their families for taking part in this study and the funding agencies that supported this work, including grants from the National Institute of Mental Health (MH74849 to IHG, MH090617 to LCFR and MH020016 to MDS), the National Alliance for Research in Schizophrenia and Affective Disorders (NARSAD; Distinguished Investigator Award to IHG and Young Investigator Award 19018 to LCFR), the National Science Foundation (Integrative Graduate Education and Research Traineeship (IGERT) 0801700 and Graduate Research Fellowship Program (GRFP) DGE-1147470 to MDS) and from the Hope for Depression Research Foundation (to IHG and LCFR).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ijdevneu.2015.07.007>.

## References

- Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.
- Belden, A.C., Barch, D.M., Oakberg, T.J., April, L.M., Harms, M.P., Botteron, K.N., Luby, J.L., 2015. Anterior insula volume and guilt: neurobehavioral markers of recurrence after early childhood major depressive disorder. *JAMA Psychiatry* 72, 40–48.
- Black, J.M., Tanaka, H., Stanley, L., Nagamine, M., Zakerani, N., Thurston, A., Kesler, S., Hulme, C., Lyytinen, H., Glover, G.H., Serrone, C., Raman, M.M., Reiss, A.L., Hoefl, F., 2012. Maternal history of reading difficulty is associated with reduced language-related gray matter in beginning readers. *Neuroimage* 59, 3021–3032.
- Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., Noll, D.C., 1997. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5, 49–62.
- Bremner, J.D., Vythilingam, M., Vermetten, E., Nazeer, A., Adil, J., Khan, S., Staib, L.H., Charney, D.S., 2002. Reduced volume of orbitofrontal cortex in major depression. *Biol. Psychiatry* 51, 273–279.
- Chen, M.C., Burley, H.W., Gotlib, I.H., 2012. Reduced sleep quality in healthy girls at risk for depression. *J. Sleep Res.* 21, 68–72.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learn.* 20 (3), 273–297.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194.
- De Martino, F., Valente, G., Staeren, G., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of functional MRI spatial patterns. *Neuroimage* 43, 44–58.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980.
- American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (4th ed., Text Revision). American Psychiatric Association, Washington, DC.
- Dickerson, B.C., Fenstermacher, E., Salat, D.H., Wolk, D.A., Maguire, R.P., Desikan, R., Pacheco, J., Quinn, B.T., Van der Kouwe, A., Greve, D.N., Blacker, D., Albert, M.S., Killiany, R.J., Fischl, B., 2008. Detection of cortical thickness correlates of cognitive performance: reliability across MRI scan sessions, scanners, and field strengths. *Neuroimage* 39, 10–18.
- Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America* 97, 11050–11055.
- Foland-Ross, L.C., Gilbert, B.L., Joormann, J., Gotlib, I.H. (in press). Neural markers of familial risk for depression: an investigation of cortical thickness abnormalities in healthy adolescent daughters of mothers with recurrent depression. *Journal of Abnormal Psychology*.
- Frodl, T.S., Koutsouleris, N., Bottlender, R., Born, C., Jäger, M., Scupin, I., Reiser, M., Möller, H.J., Meisenzahl, E.M., 2008. Depression-related variation in brain morphology over 3 years: effects of stress? *Arch. Gen. Psychiatry* 65, 1156–1165.
- Fu, C.H., Mourao-Miranda, J., Costafreda, S.G., Khanna, A., Marquand, A.F., Williams, S.C., Brammer, M.J., 2008. Pattern classification of sad facial processing: toward

- the development of neurobiological markers in depression. *Biol. Psychiatry* 63, 656–662.
- Gladstone, T.R.G., Beardslee, W.R., O'Connor, E.E., 2011. The prevention of adolescent depression. *Psychiat. Clin. North Am.* 34, 35–52.
- Gorwood, P., 2008. Neurobiological mechanisms of anhedonia. *Dialogues Clin. Neurosci.* 10, 291–299.
- Gotlib, I.H., Hamilton, J.P., Cooney, R.E., Singh, M.K., Henry, M.L., Joormann, J., 2010. Neural processing of reward and loss in girls at risk for major depression. *Arch. Gen. Psychiatry* 67, 380–387.
- Grieve, S.M., Korgaonkar, M.S., Koslow, S.H., Gordon, E., Williams, L.M., 2013. Widespread reductions in gray matter volume in depression. *Neuroimage Clin.* 3, 332–339.
- Grotegerd, D., Suslow, T., Bauer, J., Ohrmann, P., Arolt, V., Stuhmann, A., Heindel, W., Kugel, H., Dannlowski, U., 2013. Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: a pilot study. *Eur. Arch. Psychiatry Clin. Neurosci.* 263, 119–131.
- Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. *Math. Intelligencer* 27, 83–85.
- Huang, Y.M., Du, S., 2005. Weighted support vector machine for classification with uneven training class sizes. *Proceedings of 2005 International Conference on Machine Learning and Cybernetics vol. 7*, 4365–4369.
- Japkowicz, N., Stephen, S., 2002. The class imbalance problem: a systematic study. *Intelligent Data Anal.* 6, 429–449.
- Joormann, J., Talbot, L., Gotlib, I.H., 2007. Biased processing of emotional information in girls at risk for depression. *J. Abnorm. Psychol.* 116, 135.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., Ryan, N., 1997. Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 980–988.
- Kerestes, R., Davey, C.G., Stephanou, K., Whittle, S., Harrison, B.J., 2014. Functional brain imaging studies of youth depression: a systematic review. *Neuroimage Clin.* 4, 209–231.
- Kessler, R.C., Avenevoli, S., Ries Merikangas, K., 2001. Mood disorders in children and adolescents: an epidemiologic perspective. *Biol. Psychiatry* 49, 1002–1014.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S., 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 289, 3095–3105.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International Joint Conference on Artificial Intelligence*, 1137–1145.
- Koolschijn, P.C., van Haren, N.E., Lensvelt-Mulders, G.J., Hulshoff Pol, H.E., Kahn, R.S., 2009. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Hum. Brain Mapp.* 30, 3719–3735.
- Kovacs, M., 1992. *Children's Depression Inventory*, 2nd ed. Multi-Health Systems.
- Lacerda, A.L., Keshavan, M.S., Hardan, A.Y., Yorbik, O., Brambilla, P., Sassi, R.B., Nicoletti, M., Mallinger, A.G., Frank, E., Kupfer, D.J., Soares, J.C., 2004. Anatomic evaluation of the orbitofrontal cortex in major depressive disorder. *Biol. Psychiatry* 55, 353–358.
- Lamm, C., Singer, T., 2010. The role of anterior insular cortex in social emotions. *Brain Struct. Funct.* 214, 579–591.
- Lewinsohn, P.M., Clarke, G.N., Seeley, J.R., Rohde, P., 1994. Major depression in community adolescents: Age at onset, episode duration, and time to recurrence. *J. Am. Acad. Child Adolesc. Psychiatry* 33, 809–818.
- Lewinsohn, P.M., Rohde, P., Seeley, J.R., 1998. Treatment of adolescent depression: frequency of services and impact on functioning in young adulthood. *Depress. Anxiety* 7, 47–52.
- March, J.S., Parker, J.D.A., Sullivan, K., Stallings, P., Conners, C.K., 1997. The Multidimensional Anxiety Scale for Children (MASC): Factor Structure, Reliability, and Validity. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 554–565.
- March, J.S., Silva, S., Petycky, S., for, A., with, D., epression, S., tudy, T., 2004. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression. *JAMA* 292, 807–820.
- Mennon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667.
- Merikangas, K.R., He, J.P., Burstein, M., Swanson, S.A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *J. American Acad. Child Adolesc. Psychiatry*, 49, 980–989.
- Miller, C., Hamilton, J.P., Sacchet, M.D., Gotlib, I.H. (in press) A quantitative whole-brain meta-analysis of emotion processing in adolescent Major Depressive Disorder. *JAMA Psychiatry*.
- Murray, C., Lobez, A., 1996. *The Global Burden of Disease*. Harvard School of Public Health, Cambridge.
- Nakano, M., Matsuo, K., Nakashima, M., Matsubara, T., Harada, K., Egashira, K., Masaki, H., Takahashi, K., Watanabe, Y., 2014. Gray matter volume and rapid decision-making in major depressive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 48, 51–56.
- Nolen-Hoeksema, S., Hilt, L.M., 2010. Gender differences in depression. In: *Handbook of Depression*, 2nd Ed. Guilford Press, pp. 386–404.
- Orrù, G., Pettersson-Yeo, W., Marquand, A.F., Sartori, G., Mechelli, A., 2012. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci. Biobehav. Rev.* 36, 1140–1152.
- Pagliaccio, D., Luby, J.L., Luking, K.R., Belden, A.C., Barch, D.M., 2014. Brain-behavior relationships in the experience and regulation of negative emotion in healthy children: implications for risk for childhood depression. *Dev. Psychopathol.* 26, 1289–1303.
- Pakkenberg, B., Gundersen, H.J., 1997. Neocortical neuron number in humans: effect of sex and age. *J. Comp. Neurol.* 384, 312–320.
- Papmeyer, M., Giles, S., Sussmann, J.E., Kieley, S., Stewart, T., Lawrie, S.M., Whalley, H.C., McIntosh, A.M., 2014. Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. *Biol. Psychiatry*.
- Peterson, B., Warner, V., Bansal, R., Zhu, H., Hao, X., Liu, J., Durkin, K., Adams, P.B., Wickramaratne, P., Weissman, M., 2009. Cortical thinning in persons at increased familial risk for major depression. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6273–6278.
- Pizzagalli, D.A., 2014. Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annu. Rev. Clin. Psychol.* 10, 393–423.
- Qiu, L., Lui, S., Kuang, W., Huang, X., Li, J., Li, J., Zhang, J., Chen, H., Sweeney, J.A., Gong, Q., 2014. Regional increases of cortical thickness in untreated, first-episode major depressive disorder. *Transl. Psychiatry* 4, e378.
- Raudys, S.J., Jain, A.K., 1991. Small sample size effects in statistical pattern recognition: recommendations for practitioners. *IEEE Trans. Pattern Anal. Machine Intelligence* 13, 252–264.
- Rawal, A., Collishaw, S., Thapar, A., Rice, F., 2013. The risks of playing it safe: a prospective longitudinal study of response to reward in the adolescent offspring of depressed parents. *Psychol. Med.* 43, 27–38.
- Redlich, R.A., Almeida, J.J.R., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., Arolt, V., Phillips, M.L., Dannlowski, U., 2014. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry* 71, 1222–1230.
- Reynolds, S., Carrey, N., Jaworska, N., Langevin, L.M., Yang, X.-R., Macmaster, F.P., 2014. Cortical thickness in youth with major depressive disorder. *BMC Psychiatry* 14, 83.
- Rolls, E., 1996. The orbitofrontal cortex. *Philos. Trans. R. Soc. Lond.* 351, 1433–1444.
- Salat, D., Buckner, R., Snyder, A., Greve, D., Desikan, R., Busa, E., Morris, J.C., Dale, A.M., Fischl, B., 2004. Thinning of the cerebral cortex in aging. *Cereb. Cortex* 14, 721–730.
- Tao, R., Emslie, G.J., Mayes, T.L., Nakonezny, P.A., Kennard, B.D., 2010. Symptom improvement and residual symptoms during acute antidepressant treatment in pediatric major depressive disorder. *J. Child Adolesc. Psychopharmacol.* 20, 423–430.
- Van Eijndhoven, P., van Wingen, G., Katzenbauer, M., Groen, W., Tepest, R., Fernández, G., Buitelaar, J., Tendolcar, I., 2013. Paralimbic cortical thickness in first-episode depression: evidence for trait-related differences in mood regulation. *Am. J. Psychiatry* 170, 1477–1486.
- Weissman, M.M., Warner, V., Wickramaratne, P., Moreau, D., Olfson, M., 1997. Offspring of depressed parents. 10 Years later. *Arch. Gen. Psychiatry* 54, 932–940.
- Weschler, D., 1991. *The Weschler Intelligence scale for children*. Psychol. Corporation.
- Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C., 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage* 53, 1135–1146.
- Whiteford, H.A., Degenhardt, L., Rehm, J., Baxter, A.J., Ferrari, A.J., Erskine, H.E., Charlson, F.J., Norman, R.E., Flaxman, A.D., Johns, N., Burstein, R., Murray, C.J., Vos, T., 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 382, 1575–1586.
- Yang, Y., Raine, A., Colletti, P., Toga, A.W., Narr, K.L., 2009. Abnormal temporal and prefrontal cortical gray matter thinning in psychopaths. *Mol. Psychiatry* 14, 561–562, 555.