

# UCLA

## UCLA Previously Published Works

### Title

Inter and intra-hemispheric structural imaging markers predict depression relapse after electroconvulsive therapy: a multisite study

### Permalink

<https://escholarship.org/uc/item/13v073pp>

### Journal

Translational Psychiatry, 7(12)

### ISSN

2158-3188

### Authors

Wade, Benjamin SC

Sui, Jing

Hellemann, Gerhard

et al.

### Publication Date

2017

### DOI

10.1038/s41398-017-0020-7

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

ARTICLE

Open Access

# Inter and intra-hemispheric structural imaging markers predict depression relapse after electroconvulsive therapy: a multisite study

Benjamin S. C. Wade<sup>1,2</sup>, Jing Sui<sup>3,4,5</sup>, Gerhard Hellemann<sup>6</sup>, Amber M. Leaver<sup>1</sup>, Randall T. Espinoza<sup>1,2</sup>, Roger P. Woods<sup>1</sup>, Christopher C. Abbott<sup>7</sup>, Shantanu H. Joshi<sup>1</sup> and Katherine L. Narr<sup>1,2</sup>

## Abstract

Relapse of depression following treatment is high. Biomarkers predictive of an individual's relapse risk could provide earlier opportunities for prevention. Since electroconvulsive therapy (ECT) elicits robust and rapidly acting antidepressant effects, but has a >50% relapse rate, ECT presents a valuable model for determining predictors of relapse-risk. Although previous studies have associated ECT-induced changes in brain morphometry with clinical response, longer-term outcomes have not been addressed. Using structural imaging data from 42 ECT-responsive patients obtained prior to and directly following an ECT treatment index series at two independent sites (UCLA:  $n = 17$ , age =  $45.41 \pm 12.34$  years; UNM:  $n = 25$ ; age =  $65.00 \pm 8.44$ ), here we test relapse prediction within 6-months post-ECT. Random forests were used to predict subsequent relapse using singular and ratios of intra and inter-hemispheric structural imaging measures and clinical variables from pre-, post-, and pre-to-post ECT. Relapse risk was determined as a function of feature variation. Relapse was well-predicted both within site and when cohorts were pooled where top-performing models yielded balanced accuracies of 71–78%. Top predictors included cingulate isthmus asymmetry, pallidal asymmetry, the ratio of the paracentral to precentral cortical thickness and the ratio of lateral occipital to pericalcarine cortical thickness. Pooling cohorts and predicting relapse from post-treatment measures provided the best classification performances. However, classifiers trained on each age-disparate cohort were less informative for prediction in the held-out cohort. Post-treatment structural neuroimaging measures and the ratios of connected regions commonly implicated in depression pathophysiology are informative of relapse risk. Structural imaging measures may have utility for devising more personalized preventative medicine approaches.

## Introduction

Major depressive disorder (MDD) has a lifetime prevalence of 16% in the U.S., with comparable high rates in other nations<sup>1</sup>. Approximately 40% of depressed patients do not respond to standard first-line treatments<sup>2,3</sup>. Following the first major depressive episode, >50% of recovering patients suffer relapse and >15% experience

unremitting, chronic symptoms<sup>4</sup>; recurrence will occur in ~80% of patients with a history of two or more prior episodes<sup>5–7</sup>.

The high relapse and recurrence rates, usually defined as deterioration to the full disease syndrome during a period of remission or the appearance of a new episode after a longer period of recovery, respectively<sup>8</sup>, underscore the need for timely intervention. Identifying clinical, demographic, and physiological markers predictive of relapse and symptom recurrence are thus critical for prevention. Previous reports have associated incomplete recovery<sup>9</sup>, illness duration and number of previous

Correspondence: Katherine L. Narr ([narr@ucla.edu](mailto:narr@ucla.edu))

<sup>1</sup>Department of Neurology, UCLA, Ahmanson-Lovelace Brain Mapping Center, Los Angeles, USA

<sup>2</sup>Department of Psychiatry and Biobehavioral Sciences, UCLA, Los Angeles, USA  
Full list of author information is available at the end of the article

© The Author(s) 2017



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

episodes with long-term clinical outcomes<sup>8,10</sup>. Fewer studies have related neuroimaging measures to clinical outcome following treatment. Of these, the majority have focused on prediction of acute response or remission directly following treatment. Even less have attempted to predict a patient's likelihood of relapse/recurrence in the months following treatment.

Of prior studies addressing neuroimaging markers of long-term clinical outcomes following antidepressant treatment, one investigation following 30 MDD patients prospectively over 3-years reported that those with smaller hippocampal volumes and recurrent depression had worse outcomes, irrespective of continued medication status<sup>11</sup>. In a follow-up study, investigators showed that patients who remitted during the 3-year timeframe had less atrophy of the left hippocampus, left anterior cingulate (ACC), and left dorsomedial prefrontal cortex, lateralized to the left hemisphere, and bilateral dorso-lateral prefrontal cortex, than non-remitters<sup>12</sup>. Corroborating evidence from a recent study following 49 MDD patients over 5-years similarly found patients with smaller ACC volumes (though right-lateralized) prior to treatment had poorer clinical outcomes<sup>13</sup>. Further, inclusion of volumetric measures of both the right ACC and right inferior frontal cortex increased the explained variance for change in mood by 20% over inclusion of clinical and demographic measures alone. These findings support the potential utility of structural imaging measures as biomarkers of clinical outcome<sup>13</sup>.

Using naturalistic designs, others have shown that task-related brain activation within dorsolateral and medial prefrontal regions, striatum, and parietal regions are predictive of a chronic versus a more favorable clinical course of depression with up to 73% accuracy<sup>14</sup>. Interestingly, structural imaging measures were not predictive in this particular study. Another group also found that activation of the ACC and ventromedial prefrontal cortex during an emotional challenge task was strongly associated with increased relapse risk ~18-months following remission. Using a signal change threshold of  $\geq 0\%$  and regions-of-interest targeted from statistical remitter-control contrasts, relapsing patients were classified with a sensitivity and specificity of 90% and 83%, respectively<sup>15</sup>.

Notably, while these studies made important strides towards identifying biomarkers of clinical outcome, with few exceptions<sup>14</sup>, these investigations tested for post-hoc associations rather than the predictive utility of imaging biomarkers in a cross-validation framework<sup>11–13,15</sup>. Further, prospective studies have followed subjects within a naturalistic setting where patients have received several possible treatment regimens<sup>13,14</sup>. Generalization across treatments is important, but since different treatment strategies may have different response/relapse trajectories, different degrees of compliance and differentially affect a

patient's neurobiology, focus on a single treatment modality may be beneficial for targeting neural predictors of relapse and symptom recurrence in MDD.

Here, we focus on predicting relapse/recurrence within 6-months following treatment with electroconvulsive therapy (ECT), which is a well-established treatment typically reserved for patients with severe treatment-resistant depression. ECT works more quickly (response can occur in 2–4 weeks) and has higher remission rates than other standard therapeutic approaches<sup>16</sup>. However, relapse/recurrence rates of depression after ECT are similar to those of other antidepressant treatments. Following ECT, relapse occurs in ~50% of patients with most relapsing within the first 6-months<sup>17</sup>. The fast acting and robust clinical effects of ECT together with relapse and recurrence risks comparable to other treatments make this treatment ideal for determining whether variations in brain morphology are predictive of individual clinical outcome.

Selecting 42 patients who initially responded to ECT from two studies conducted at independent sites, the University of California, Los Angeles (UCLA) and the University of New Mexico (UNM), we developed a random forest (RF) classifier to identify structural neuroimaging, clinical, and demographic factors predictive of symptom recurrence/relapse. Imaging measures were derived from pre-treatment, post-treatment, and pre-post treatment change in subcortical volumes and regional cortical thickness and ratios between these measures. The motivation for the latter included the following: (i) ratios for cortical thickness/volumes are normalized within subject, (ii) they capture effects of asymmetry since the ratios of homologous regions are included, and (iii) they represent structural relationships between spatially diffuse regions in a spirit similar to modeling structural networks. Because ECT was administered unilaterally in the majority of patients, we hypothesized that measures of asymmetry would relate to relapse risk.

The identification of a set of biomarkers informing a patient's probability of symptom recurrence following treatment is of high translational value. Clinicians could refine maintenance strategies for at-risk individuals to prevent relapse, which could offset the disproportionately high cost of managing recurrent major depression.

## Patients and Methods

### Participants

Patients experiencing a DSM-IV defined major depressive episode and eligible to receive ECT were recruited from UCLA ( $N = 42$ ) and UNM ( $N = 40$ ). All patients received structural MRI scans and mood evaluations 24 h prior to ECT (pretreatment) and within a week of completing ECT index (post-treatment). Patient mood was again assessed approximately 6-months following index (follow-up). The Hamilton Depression Rating Scale

(HAM-D-17)<sup>18</sup> tracked symptomology at each time point. A patient was defined to have relapsed if (i) their HAM-D reduced by  $\geq 50\%$  over ECT index indicating therapeutic response and (ii) their long-term follow-up HAM-D score was  $\geq 17$ . The threshold for determining response was selected since it the most commonly used definition of clinically meaningful response<sup>19</sup>. Definitions of relapse following initial response are less consistent. Since the current study (a) included patients with severe and treatment resistant depression, (b) the criterion used for response still allowed for some residual symptoms, and (c) HAM-D scores of  $\geq 17$  are considered a cut-off for separating mild<sup>8–16</sup> from moderate<sup>17–23</sup> and severe depression ( $\geq 24$ )<sup>20</sup>, a HAM-D score of  $\geq 17$  at final follow-up was used to define relapse. A total of 17 patients at UCLA (10 females, mean age =  $45.41 \pm 12.34$ ) and 25 at UNM (18 females, mean age =  $65 \pm 8.44$ ) were defined as ECT responders and included in the study. Following the ECT index, patients continued on naturalistic course of maintenance therapy. Ten UCLA patients and 11 at UNM received maintenance or continuation (m-/c-ECT), respectively. Within 6 months following ECT index 6 patients (35%) relapsed at UCLA while 13 (52%) relapsed at UNM.

All patients had experienced two or more earlier major depressive episodes and failed to respond to at least two prior adequate medication trials in the index episode. Exclusionary criteria included first-episode depression, diagnosed neurological or neurodegenerative disorder, any head injury with loss of consciousness over 5 min, comorbid psychiatric conditions such as schizophrenia or schizoaffective disorder, current drug or alcohol abuse (excluding nicotine), and MRI contraindications. Bipolar disorder was exclusionary at UNM but not at UCLA, though mania in the index episode was exclusionary at UCLA. UCLA patients were excluded if the age of depression onset was over 50 years. UCLA patients were tapered off of antidepressants and benzodiazepines in preparation for ECT and were completely free of medication for at least 48–72 h before enrollment and ECT treatment. The UNM cohort was not tapered off of medication before ECT. All participants provided written informed consent as approved by the UCLA or UNM Institutional Review Board.

### MR acquisition

At UCLA, high-resolution motion-corrected multi-echo T1-weighted MPRAGE structural brain images<sup>21</sup> were acquired on a Siemens 3 T Allegra system (Erlangen, Germany) for all subjects and time points (TEs/TR = 1.74, 3.6, 5.46, 7.32/2530 ms, TI = 1260 ms, flip angle =  $7^\circ$ , voxel resolution =  $1.3 \times 1 \times 1 \text{ mm}^3$ ). Patients at UNM were scanned on a 3-Tesla Siemens Trio scanner with a similar multi-echo T1-weighted MPRAGE (TR = 2.53 s

(s), TE = 1.64, 3.5, 5.32, 7.22, 9.08 ms, TI = 1.20 s, flip angle =  $7^\circ$ , number of excitations = 1, and voxel resolution =  $1 \times 1 \times 1 \text{ mm}^3$ ).

### Image preprocessing

Validated FreeSurfer<sup>22</sup> workflows, including removal of non-brain tissue, intensity normalization and automated volumetric parcellation based on probabilistic information from manually labeled training sets, were used for whole brain cortical (Desikan Killiany atlas-based parcellations<sup>23</sup>) and subcortical segmentation. Each segmentation was visually inspected to ensure its quality.

### Candidate features

Clinical, demographic and structural neuroimaging features of each participant were included as candidate features for the RF classifier. Demographic information included age and sex; clinical information was comprised of pre- and post-treatment HAM-D-17 scores, number of ECT sessions received, and electrode placement (right unilateral or bilateral). Maintenance therapy status (m-/c-ECT) was not a predictor. Imaging features included volumes of the accumbens, amygdala, caudate, hippocampus, pallidum, putamen, thalamus, and lateral ventricles and the mean thickness of 34 homologous cortical regions.

In addition to individual measures of subcortical volumes and cortical thickness we included pairwise ratios of each subcortical volume to every other subcortical volume and likewise for cortical thickness. These ratios have several important benefits. First, they are normalized within subject and thus robust to confounds of age- and sex-related group differences. Ratio measures additionally capture the asymmetry of homologous regions that may reflect underlying neurobiological properties related to clinical outcome. Second, the distribution of ECT electrode placements largely determines the electrical field distribution within the brain and the distribution of ECT's direct effects. Since unilateral electrode placement is mostly applied at UCLA and UNM, there are likely to be highly lateralized effects, which may also relate to clinical outcome.

### Predictive modeling

We used supervised RFs<sup>24</sup> to classify relapse status based on either pre- or post-treatment features or changes in these features over treatment. RFs were used due to their robustness against overfitting in high-dimensions, minimal tuning parameters, and overall competitive performance relative to other state-of-the-art methods<sup>25</sup>. Alternative classifiers were not compared to RFs in order to mitigate type 1 errors. We considered models built from participants exclusively at UCLA or UNM and observations combined across sites. Models fit to a single

cohort were additionally assessed by predicting relapse across site.

We implemented leave-one-out cross-validation (LOO-CV) to validate our classifier. Within each iteration of LOO-CV, the set of  $N-1$  participants used to train the model was partitioned into 10 nested folds to further randomize the feature selection process which consisted of multiple steps. We first held out one of the 10 folds. Second, features in the remaining folds were subjected to a collinearity filter in which we identified pairs of features correlated above a threshold,  $|r|$ . The feature in this pair with the highest correlation with the remaining features in the dataset was excluded. The optimal threshold,  $|r|$ , was found using a grid search over the space  $r \in \{0.1, 0.2, \dots, 1.0\}$ . The remaining features were then passed to a recursive feature elimination<sup>26</sup> (RFE) algorithm where the internal classifier was a RF composed of 1000 trees and the feature set maximizing classification specificity was selected.

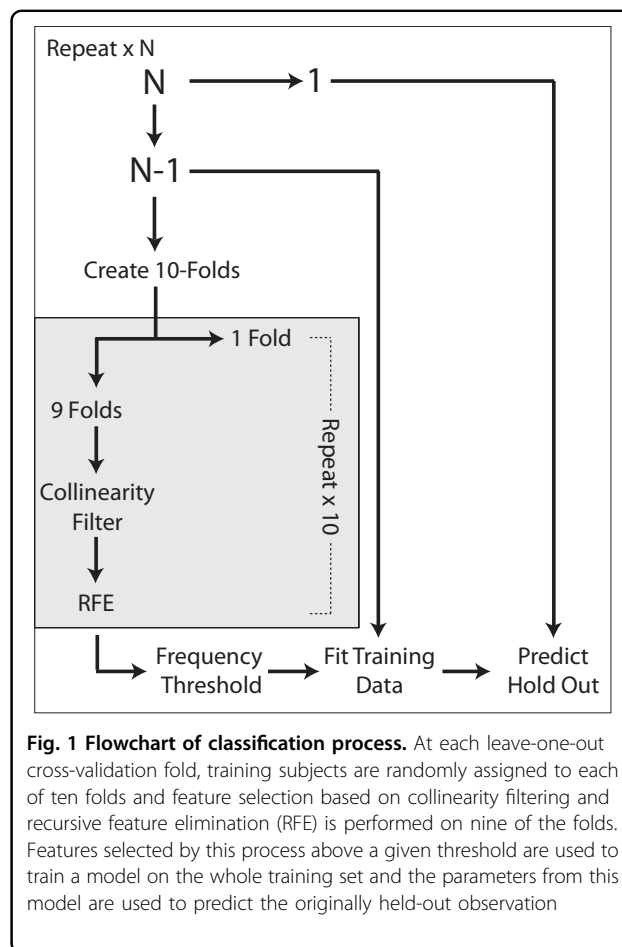
After this process was repeated 10 times, we evaluated the frequency with which each feature was selected by RFE. We retained features selected above the  $\rho$ -th quantile of frequencies where  $\rho \in \{0.1, 0.2, \dots, 1.0\}$ . Finally, the features identified by this process were used to fit a RF to the entire  $N-1$  set of training observations. RF-specific tuning parameters are outlined in Supplementary Methods. The final RF predicted the originally held-out observation (Fig. 1). A grid search was performed over the space of parameters  $r$  and  $\rho$  for each location and time-point specific model. Models constructed within-site were additionally used to predict the patient relapse across site as further validation. All RFs were implemented using the *caret*<sup>27</sup> package in R version 3.3.2<sup>28</sup>.

Performance measures in binary classification with imbalanced class proportions are biased towards the more prevalent class. To avoid reporting optimistic performances we report the balanced accuracy (BA)<sup>29</sup> defined as the arithmetic mean of sensitivity and specificity. We additionally compare the BA to the baseline detection rate (BDR), which is the accuracy obtained by hypothetically assigning each unknown patient to the most prevalent class. The BDR is naturally  $>50\%$  in problems with imbalanced classes.

Two post-hoc analyses were also conducted (see Supplementary Methods and Results). First we explored our model's performance as a function of classifying only patients for which the classifier had more than a particular level of confidence, known as classification with a rejection option<sup>30</sup>. We secondly repeated our entire analysis using 10 repeated 10-fold cross-validation to assess the classifier's robustness across cross-validation schemes.

#### Code availability

Code developed for these analyses are available upon request.



**Fig. 1 Flowchart of classification process.** At each leave-one-out cross-validation fold, training subjects are randomly assigned to each of ten folds and feature selection based on collinearity filtering and recursive feature elimination (RFE) is performed on nine of the folds. Features selected by this process above a given threshold are used to train a model on the whole training set and the parameters from this model are used to predict the originally held-out observation

## Results

### Demographic and clinical measures

Table 1 shows patient demographic and clinical characteristics. The UNM cohort was significantly older than the UCLA cohort. Post-treatment HAM-D scores were significantly lower at UNM. However, the cohorts did not differ significantly by sex, average number of ECT sessions, proportion of patients receiving RUL electrode placement, pretreatment HAM-D or the proportion relapsers.

All relapsing patients were female at UCLA; 10 of the 13 relapsing patients were female at UNM. This proportion did not differ significantly from non-relapsers. Across both sites, relapsers did not differ significantly from non-relapsers in terms of pre- or post-treatment HAM-D scores, number of ECT sessions received, proportion of patients receiving right unilateral electrode placement, age, or m-/c-ECT status.

### Prediction from pre-treatment measures

Here we detail the highest performing model from each location provided their BAs are above the BDR of the

**Table 1 Demographic and clinical characteristics**

	UCLA, N = 17	UNM, N = 25	t or $\chi^2$ (p-value)
Age, mean (SD), years	45.41 (12.34)	65.00 (8.44)	$p = 5.34e-06$
Sex, (M/F)	7/10	7/18	$p = 0.57$
# of ECT index sessions, mean (SD)	10.11 (2.31)	10.36 (3.05)	$p = 0.77$
# of RUL arrangements, (RUL/BT)	16/1	21/4	$p = 0.61$
Pretreatment HAM-D-17, mean (SD)	27.29 (5.24)	26.44 (7.26)	$p = 0.66$
Post-treatment HAM-D-17, mean (SD)	6.82 (3.28)	3.12 (3.07)	$p = 0.00082$
# Relapsers/Non-relapsers	6/11	13/12	$p = 0.45$
# Unipolar/Bipolar	14/3	25/0	$p = 0.11$
	UCLA Relapsers, N = 6	UCLA Non-relapsers N = 11	Significance
Age, mean (SD), years	43.16 (13.89)	46.63 (11.94)	$p = 0.61$
Sex, (M/F)	0/6	7/4	$p = 0.04$
# of ECT index sessions, mean (SD)	10.00 (1.67)	10.18 (2.67)	$p = 0.86$
# of RUL arrangements, (RUL/BT)	6/0	10/1	$p = 1$
Pretreatment HAM-D-17, mean (SD)	28.83 (4.79)	26.45 (5.50)	$p = 0.37$
Post-treatment HAM-D-17, mean (SD)	6.66 (2.65)	6.90 (3.70)	$p = 0.87$
# Maintenance ECT, (Yes/No)	5/1	5/6	$p = 0.30$
	UNM Relapsers, N = 13	UNM Non-relapsers N = 12	Significance
Age, mean (SD), years	65.15 (9.76)	64.83 (7.17)	$p = 0.92$
Sex, (M/F)	3/10	4/8	$p = 0.90$
# of ECT index sessions, mean (SD)	10.53 (3.38)	10.16 (2.79)	$p = 0.76$
# of RUL arrangements, (RUL/BT)	10/3	11/1	$p = 0.64$
Pretreatment HAM-D-17, mean (SD)	25.92 (6.10)	27.00 (8.60)	$p = 0.72$
Post-treatment HAM-D-17, mean (SD)	3.23 (3.46)	3.00 (2.73)	$p = 0.85$
# Continuation ECT, (Yes/No)	6/7	5/7	$p = 1$

UCLA University of California, Los Angeles, UNM University of New Mexico, ECT Electroconvulsive therapy, RUL Right unilateral electrode placement, BT Bi-temporal, HAM-D Hamilton depression rating scale.

cohort used to train the model. Classifier performances from pre-, post-, and change during ECT series for individual and combined sites are presented in Table 2. Regions important to each classifier are illustrated in Fig. 2. In Fig. 3 we plot the BA obtained across all parameterizations of the models explored in the grid search and compare this distribution to the model's BDR.

Relapse at UCLA was well-predicted from pre-treatment measures with a BA of 78%; well above the 68% BDR. Cingulate isthmus thickness asymmetry was the most important feature, followed by the ratio of left superior frontal to left rostral middle frontal cortical thickness. In post-hoc examination of these measures, relapsing patients were shown to have numerically smaller, though not statistically different, right to left hemisphere cingulate isthmus thickness ratios than non-

relapsers. The asymmetry of the cingulate isthmus trended towards a significant association with sex ( $t = 2.075$ ,  $p = 0.056$ , Supplementary Fig. 4). No statistically significant post-hoc relapse/non-relapser difference for the left superior frontal to left rostral middle frontal cortical thickness ratio was observed. While several UCLA-based models exceeded the BDR, many did not, suggesting some degree of model instability. Supplementary Table 1 further outlines important features that were selected by each classifier in at least 50% of LOO-CV folds.

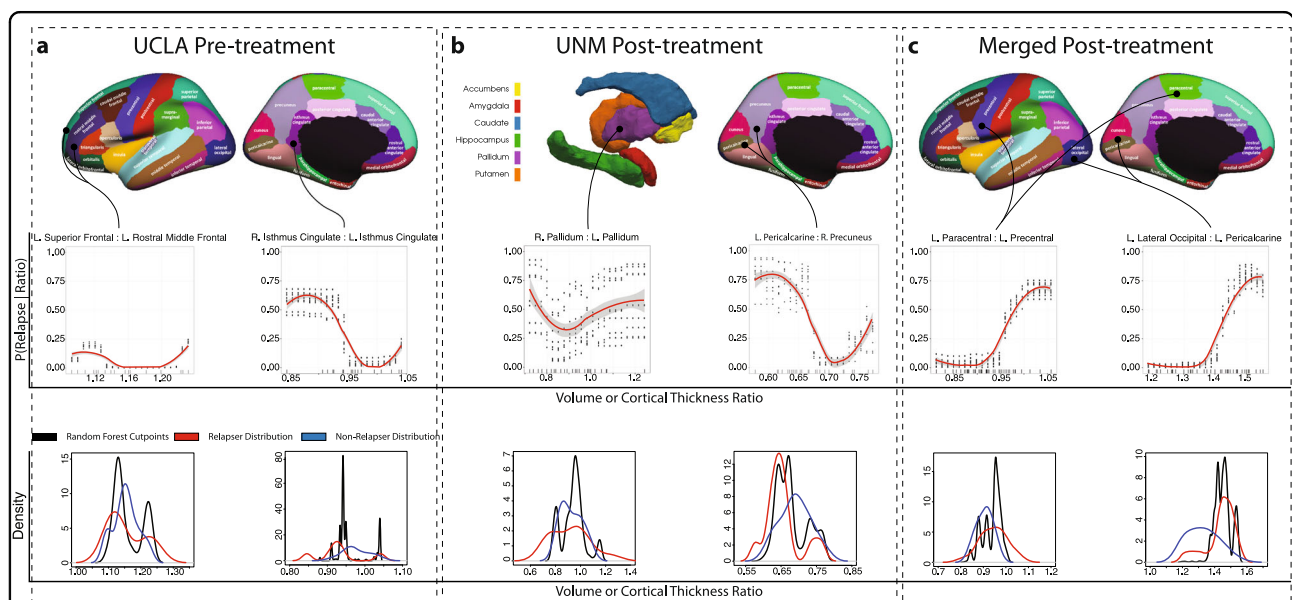
Using RFs fitted to 10 randomly resampled sets of observations we modeled the predicted probability of a simulated patient's relapse given an observed range of these two ratios. Each simulated patient was assigned the within-cohort average of each feature except for the ratio measure to be evaluated, which itself took on a series of 20



**Table 2 Classifier performance summaries: highest-performance models by site and time point**

Location	Time point	Accuracy	Balanced Accuracy	Sensitivity	Specificity	PPV	NPV	BDR (Balanced Accuracy)
UCLA	T1	82%	78%	90%	66%	83%	80%	68%(10)
	T2	70%	62%	90%	33%	71%	66%	68% (-6)
	TΔ	76%	66%	100%	33%	73%	100%	68 (-2)
UNM	T1	60%	60%	66%	53%	57%	63%	52% (8)
	T2	72%	71%	66%	76%	72%	71%	52% (19)
	TΔ	52%	52%	66%	38%	50%	55%	52 (0)
Merged	T1	54%	53%	65%	42%	57%	50%	54% (-1)
	T2	76%	76%	73%	78%	80%	71%	54% (22)
	TΔ	59%	59%	60%	57%	63%	55%	54% (5)

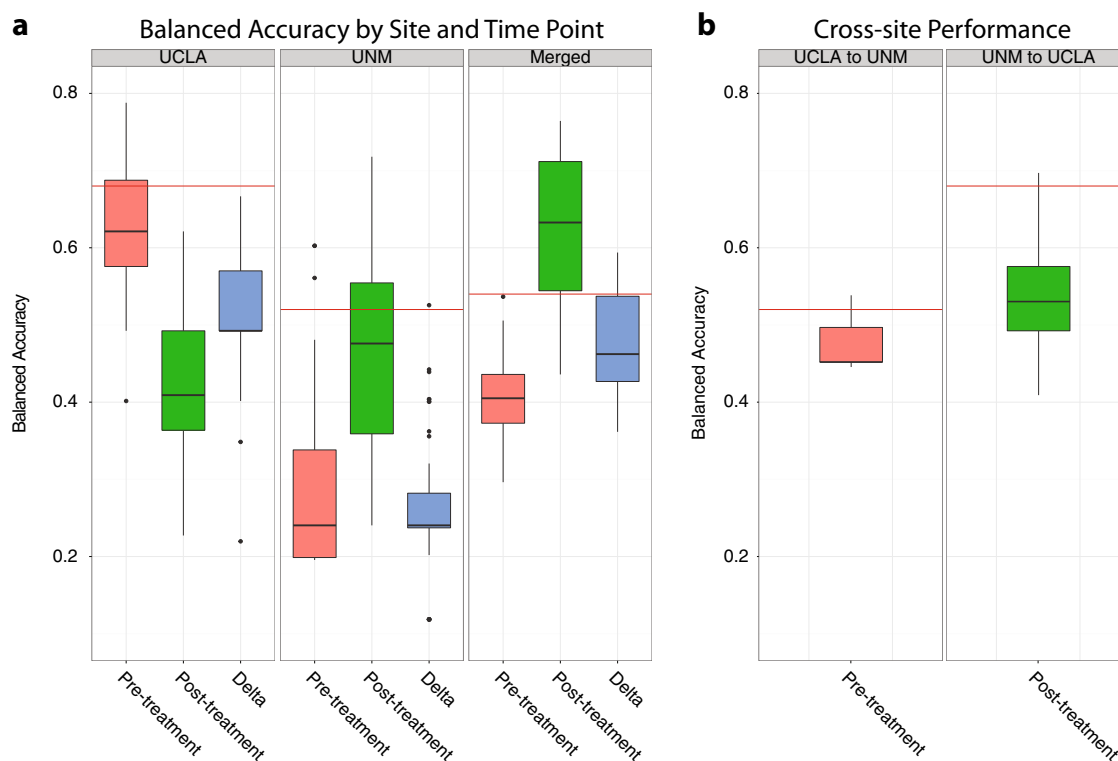
T1 Time point 1 (pretreatment), T2 Time point 2 (post-treatment), TΔ features derived from the change in values over ECT index, i.e., T2 - T1, PPV positive predictive value, NPV negative predictive value, BDR baseline detection rate. 'Positive' class is non-relapse. Balanced accuracy-BDR is a proxy for the improvement of the classifier above baseline chance levels of classification.



**Fig. 2 Regions predictive of relapse and their relationship to the predicted probability of relapse.** Top row illustrates anatomical locations of cortical and subcortical regions most important to relapse prediction. The middle row indicates the posterior probability of individual relapse over an observed range (minimum to maximum in 20 even increments) of region ratios locally averaged across 10 bootstrapped resamples of the data set and refitted to the derived classifier. A non-parametric LOESS model was fit to the predicted responses. Points about each line indicate predicted probabilities from each resample while rugs of each plot indicate the density of observed values in the whole sample. The bottom row illustrates corresponding distributions of random forest decision points (black) for these regions across underlying 1000 classification trees in the determination of relapse status. These are compared to distributions of these regions for relapsing (red) and non-relapsing (blue) patients

evenly spaced values between the observed minimum and maximum of the cohort. Predictions of the simulated patient's outcome from each of the 10 RFs were estimated using a nonparametric LOESS model (Fig. 2a). Patients with smaller right to left cingulate isthmus ratios had lower predicted probabilities of relapse. Perturbations of the left superior frontal to left rostral middle frontal ratio

exerted less influence on the posterior probability of relapse after accounting for this measure. Characteristics of the distribution of RF split points for the most important features are reported in Supplementary Results. We assessed the generalizability of all UCLA pretreatment model parameterizations that yielded a BA greater than the UCLA BDR by attempting to predict relapse at



**Fig. 3** Distributions of balanced accuracies across all classifier parameterizations. **a** Performances of classifiers trained and tested within a cohort. Horizontal red lines indicate the site-specific baseline detection rate. **b** Distributions of balanced accuracies achieved in cross-site predictions. All models exceeding their respective baseline detection rates within site for UCLA at pre-treatment (**b**, left) and UNM post-treatment (**b**, right) were used to predict relapse in the independent site. These performances are compared to the independent site's respective baseline detection rate shown in the horizontal red line

UNM from pretreatment measures using the models trained on UCLA data (see Supplementary methods). Figure 3b illustrates the distribution of these cross-site prediction performances. The mean BA obtained by predicting relapse at UNM from UCLA models was 47% (SD = 0.03%, range = 44–53%), less than the UNM BDR of 52% and the averages of the two region ratios were not consistently larger or smaller by relapse status across sites (Supplementary Fig. 3).

Using the UNM cohort to train our classifier yielded a BA of 60%, which is slightly above the 52% UNM BDR. In addition, the highest-performing RF obtained by merging the UCLA and UNM cohorts performed at near-chance levels and below the BDR (BA = 53%, BDR = 54%).

#### Prediction from post-treatment measures

Prediction of relapse at UCLA using post-treatment measures resulted in a 62% BA, below the 68% BDR. In contrast, post-treatment measures were the best predictors of relapse at UNM yielding a BA = 71%, well above the 52% BDR. Here, the right to left pallidum volume ratio was the most discriminative feature.

Groupwise means of this feature did not differ ( $t = -0.21$ ,  $p = 0.82$ ). The second most informative feature was left pericalcarine to right precuneus cortical thickness ratio, which tended (non-significantly) to be larger among non-relapsing patients ( $t = -1.94$ ;  $p = 0.06$ ). Using the fitted RF to predict simulated patients, we observed that predicted relapse was lowest when the right to left pallidum volume ratio was  $\sim 0.9$  and the left pericalcarine to right precuneus ratio was  $\sim 0.7$ , but increased steadily as these ratios deviated from these points (Fig. 2b).

RFs with BAs above the UNM BDR generally performed poorly when used to predict relapse at UCLA from post-treatment measures attaining an average BA = 54% (SD = 0.06, range = 40–69%) with most falling below the UCLA BDR of 68% (see Fig. 3b). Prediction of relapse at UCLA using post-treatment measures resulted in a BA of only 62%, below the BDR.

Classification using the combined cohorts resulted in a BA of 76% using post-treatment measures. The left lateral occipital lobe to left pericalcarine gyrus ( $t = 3.75$ ,  $p = 0.0005$ ) and left paracentral gyrus to left precentral gyrus ( $t = 2.46$ ,  $p = 0.019$ ) ratios were most informative with



both ratios being significantly larger in relapsers. Patients with larger ratios had greater a predicted probability of relapsing (Fig. 2c). Interestingly, the majority of this model's parameterizations were well above the BDR suggesting it is relatively stable across parameterizations (Fig. 3).

### Prediction from measures of change

When examining pre to post ECT-related change in structural imaging markers, both sites did not show predictive utility above their respective BDRs (UCLA = 66% and UNM = 52%). Merging the two sites resulted in a BA of 59% only slightly above the BDR of 54%.

### Discussion

To our knowledge this is the first study to predict relapse/recurrence of depression following ECT where findings are expected to be relevant to prediction of other antidepressant treatments. We used advanced classification algorithms and novel feature representations including structural neuroimaging, demographic, and clinical measures to identify biomarkers of long-term clinical outcome. Our investigation extends beyond many classification-based studies by further interrogating our fitted models to understand the relationships between salient features and a patient's relapse risk. This is an important step in the application of machine learning to data with clinical implications since a black-box algorithm yield less information about the driving mechanisms. Using these tools, we demonstrated that ratios and asymmetries of particular cortical and subcortical brain regions implicated in the pathophysiology of depression present promising biomarkers for prognosis of symptom recurrence and/or relapse.

### Significance of intra and interhemispheric regional ratios

Remarkably, using an entirely data-driven approach, we identified that the intra and interhemispheric ratios of homologous or proximal regions present biomarkers of symptom recurrence/relapse in MDD. The validity of these findings is bolstered by the observation that the selected ratios do not appear to be randomly distributed throughout the brain as we would expect if we were merely detecting noise. Instead, the constituents of these ratios were either homologous or neighboring regions.

A substantial body of evidence suggests asymmetrical neural representation of emotional control and processing<sup>31,32</sup> and structural and functional imaging findings commonly report lateralized effects in MDD<sup>33–35</sup>. Several previous studies have related hemispheric asymmetry to clinical outcome following antidepressant treatment. For example, differential patterns of asymmetric functional connectivity have been noted between rTMS responders and non-responders<sup>36</sup>. The intracranial distribution of the

electric field induced by ECT, and therefore the set of structures most affected, is shown to be impacted by electrode placement<sup>37</sup>. Since both our sites used predominantly right unilateral electrode placement, regional asymmetries and/or ratios observed post-treatment may affect clinical outcome.

### Regions influencing relapse risk

The highest performing model was obtained when the cohorts were combined. Nonetheless within-site models warrant discussion due to the significant age differences between the cohorts. ECT has been established to have greater clinical benefits in older individuals<sup>38</sup>, which suggests that changes in neural integrity with age might modulate the antidepressant mechanisms of ECT. Thus differential sets of important features across site may highlight age-related predictors of relapse risk.

Within the UCLA cohort we observed that the hemispheric balance of cingulate isthmus thickness and the ratio of left superior to left rostral middle frontal cortical thickness were highly predictive of relapse from pre-treatment ECT measures. Cingulate isthmus asymmetry was additionally associated with sex (Supplementary Fig. 4), which suggests sex is a potentially informative predictor of relapse. Women are known to be at a two-fold increased lifetime risk for MDD compared to men<sup>39</sup>. This trend is reflected both by the higher proportions of women and the higher relapse rates observed in women at both our study sites. Structural and functional abnormalities of the cingulate are repeatedly implicated in depression<sup>40,41</sup>. Notably, prior studies assessing imaging markers of longer-term outcomes have shown links with pre-treatment cingulate volumes<sup>12,13</sup>. This region forms a key part of the limbic system involved in memory and other complex cognitive functions<sup>42</sup>. Our UCLA-based model also identified the ratio of the left superior to left rostral middle frontal cortical thickness as important. This feature may point to altered reward and mood regulation circuitry, which include dorsolateral prefrontal nodes. The involvement of these regions is consistent with widespread reports of fronto-striatal-limbic network disruptions in MDD<sup>43–45</sup>.

Post-treatment measures best informed relapse prediction at UNM. The asymmetry of pallidum volume and the left pericalcarine to right precuneus cortical thickness ratio was critical to prediction. The pericalcarine and precuneus are cortical association regions in close proximity and reciprocally connected to the retrosplenial cingulate observed in the UCLA cohort. Notably, the precuneus forms part of the default mode network consistently implicated in MDD<sup>46</sup>. In a sample overlapping with this study, our group recently reported on differential ECT-induced morphological changes in the pallidum between responding and non-responding MDD patients

following ECT<sup>47</sup>. The current findings suggest these regions are also relevant to relapse. Structural pallidal abnormalities are repeatedly implicated in MDD pathology and again point to disruptions in fronto-striatal-limbic circuitry though at different network nodes<sup>48,49</sup>. Further, depressive symptomatology including amotivation, anhedonia, apathy, and rumination are often linked to abnormalities of the ventral striatum and pallidum<sup>50–52</sup>, suggesting a potential functional correlate.

Within-site classifiers were additionally tested by attempting to predict relapse at the held-out site. These classifiers' BAs were below the independent sites' BDRs. These may result from systematic differences across the two cohorts. Principally, UNM patients were significantly older than UCLA patients and only the UCLA cohort was tapered from medications during the acute phase of ECT treatment. Age-related neurodegeneration well-established<sup>53,54</sup> and interactions with depression may present confounds for classifiers trained on patients with differing ages. Further, late onset depression may have a different etiologies such as vascular contributions from cardiovascular disease and hypertension<sup>55,56</sup> that could predispose geriatric patients to depressive symptoms. Illness duration and lifetime number of depressive episodes are also likely to be greater in older individuals may impact symptom recurrence<sup>8</sup>. Given these systematic differences, future work will validate these models on well-matched independent cohorts.

The highest performing model, in terms of BA above its BDR, was obtained using post-treatment information from the combined cohorts resulting in BAs consistently above the BDR across nearly every model parameterization, whereas site-specific classifiers inconsistently outperformed the BDR. Here, the ratios of two pairs of ipsilateral and neighboring left hemisphere cortical regions were most informative in relapse prediction: the paracentral to precentral cortical thickness ratio and the lateral occipital to pericalcarine cortical thickness ratio. Paracentral thickness has been linked with impulsivity in MDD<sup>57</sup> and emotional perception and interpretation<sup>58</sup>. A recent review of 10 studies surveying 329 first-episode patients and 340 healthy controls also reported a consistent increase in the left paracentral lobe volume<sup>43</sup>. Several studies have also reported change in motor circuits in relation to MDD<sup>59</sup>. Studies using transcranial magnetic stimulation<sup>60</sup> and ECT<sup>61</sup> have specifically shown significant changes in left lateralized motor-related cortical excitability, suggesting a mechanism of treatment response, which may also impact relapse.

The second highly predictive feature in the pooled classifier was the left lateral occipital to left pericalcarine cortical thickness ratio. White matter microstructural, resting state connectivity, and volumetric abnormalities in

MDD have been noted in these regions<sup>62–65</sup>. In a study of 65 patients with recurrent MDD and 65 matched controls, Na et al. identified significantly thinner lateral occipital lobes in MDD patients versus controls<sup>66</sup>. The pericalcarine is arguably less studied or implicated in MDD. However, prior studies have shown changes in gamma-aminobutyric acid within this region in relation to different antidepressant treatments, including ECT<sup>67–69</sup>.

In the context of prior reports, the constituents of the regional ratios informing our classifiers have plausible biological grounding. Depression is considered a brain-network disorder<sup>40</sup>, thus structural abnormalities are expected and reported to occur in spatially diffuse regions. As such, a data-driven approach capturing the relative thickness/volume relationships of spatially distinct regions is valuable. As a brain-network disorder, ECT-induced neural plasticity of these spatially diffuse regions is a plausible mechanism by which ECT may mediate relapse. A body of pre-clinical and neuroimaging data suggest that changes in neural plasticity may contribute to the antidepressant effects of ECT<sup>70,71</sup>. Symptom recurrence may likewise relate to neuroplastic processes which vary across brain regions and may return to pre- or post-ECT homeostasis to predict future recurrence. Though this interpretation remains speculative, our results support that structural variations in particular brain features before and after treatment impact future relapse.

It is noteworthy that post-treatment measures yielded more accurate predictions for both UNM and the merged cohorts than pre-treatment or change measures. One plausible reason for this is that changes in brain morphology induced by ECT are highly related to relapse and this information is unavailable at pre-treatment. However, we might expect change measures to be more informative than either time point in isolation. But, from a statistical perspective, change measures include noise from both time points, which is not proportionally offset by differing signals from pre- and post-treatment measures.

### Limitations

While this is the first study to attempt prediction of depression relapse following ECT, there are important limitations. It was only possible to determine predictors of relapse in patients initially showing treatment response where the UCLA and UNM cohorts consisted of 17 and 25 ECT responders, respectively. As models using the combined cohorts consistently outperformed the BDR, while models from individual sites did not, within-site models were likely underpowered. To maximize the number of participants used to train our classifiers, we used leave-one-out cross-validation. A related limitation is that since the highest performing model was built using the merged cohorts we cannot evaluate its generalizability.

In addition, several patients transitioned from RUL to BL electrode placement during ECT index. Because BL arrangement is associated with higher remission/response rates it is also possibly associated with differential relapse rates relative to RUL. However, only a minority of patients were transitioned: Four at UNM (three of whom relapsed) and one at UCLA (a non-relapser). The small number of transitioned patients precludes statistical associations of relapse and transition status.

Differences in medication status by site is noteworthy. The UCLA cohort was tapered off of psychotropic medication before ECT unlike UNM patients. Several studies have noted that ECT increases the permeability of the blood-brain barrier<sup>72,73</sup> promoting interactions between ECT and medication. Such an interaction could augment clinical outcomes between sites.

The use of m-/c-ECT is also potentially meaningful. Although we observed no association between m-/c-ECT status and relapse rates in either site we did not include m-/c-ECT status as a predictor due to its ostensible lack of influence and because it is an unknown factor at each time point in practice.

Though the poor generalizability of the classifiers trained within site is possibly due to overfitting and/or small sample sizes, biological differences between elderly and non-elderly patients may be relevant. Since the highest performing models were derived from pooling the cohorts we anticipate improved generalizability with larger samples. Further, we acknowledge that our definition of relapse as an initially responsive patient (i.e., reduction of HAM-D by  $\geq 50\%$ ) and a subsequent return to a HAM-D  $\geq 17$  is one of several possible definitions, to avoid a potential increase in type 1 errors we did not attempt classification of relapse according to other definitions.

A further criticism is the use of a grid-search over a large parameter space whereby one may expect to identify models exceeding the BDR merely by chance. However, rather than viewing each parameterization of the model as a separate model, the identification of a set of parameters maximizing the model's performance can simply be viewed as the optimization of a single model. If, however, each reparameterization of the model is viewed as an independent model, the proportion of parameterizations yielding performances above chance can be evaluated statistically. This latter view is explored in [Supplementary Results Section](#).

## Conclusions

Though much work remains to develop models with cross-site predictive value, our current findings are novel and add important new leads. For each site we identified models that performed well above their respective BDRs. Currently, post-treatment measures appear most predictive. Models based on pretreatment measures might be

useful for clinical decisions regarding patient exposure to the unwanted side-effects of ECT if benefits are not enduring. However, post-treatment models, which yielded BAs between 71 and 78%, may contribute towards more targeted monitoring of patients at elevated risk of relapse allowing for more timely prevention strategies.

## Acknowledgements

This work is supported in part by the National Institute of Mental Health (MH092301, MH10008 and MH102743 to UCLA investigators) and the National High Tech Program (863, No. 2015AA020513 to SJ).

## Author details

<sup>1</sup>Department of Neurology, UCLA, Ahmanson-Lovelace Brain Mapping Center, Los Angeles, USA. <sup>2</sup>Department of Psychiatry and Biobehavioral Sciences, UCLA, Los Angeles, USA. <sup>3</sup>The Mind Research Network and Lovelace Biomedical and Environmental Research Institute, Albuquerque, NM, USA. <sup>4</sup>Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China. <sup>5</sup>Chinese Academy of Sciences Center for Excellence in Brain Science, Institute of Automation, Beijing, China. <sup>6</sup>Department of Psychiatry and Biobehavioral Sciences, Semel Institute, UCLA, Los Angeles, USA. <sup>7</sup>Department of Psychiatry, University of New Mexico, Albuquerque, USA

## Competing interests

The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary information

The online version of this article (doi:[10.1038/s41398-017-0020-7](https://doi.org/10.1038/s41398-017-0020-7)) contains supplementary material.

Received: 26 July 2017 Accepted: 30 July 2017

Published online: 08 December 2017

## References

- Kessler, R. C. & Bromet, E. J. The epidemiology of depression across cultures. *Annu. Rev. Public Health* **34**, 119–138 (2013).
- Trivedi, M. H. et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR\*D: implications for clinical practice. *Am. J. Psychiatry* **163**, 28–40 (2006).
- McGrath, P. J. et al. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR\*D report. *Am. J. Psychiatry* **163**, 1531–1541 (2006).
- Eaton, W. W. et al. Population-based study of first onset and chronicity in major depressive disorder. *Arch. Gen. Psychiatry* **65**, 513–520 (2008).
- Fava, M. et al. A comparison of mirtazapine and nortriptyline following two consecutive failed medication treatments for depressed outpatients: a STAR\*D report. *Am. J. Psychiatry* **163**, 1161–1172 (2006).
- Greden, J. F. The burden of disease for treatment-resistant depression. *J. Clin. Psychiatry* **62**, 26–31 (2001).
- Nemeroff, C. B. Prevalence and management of treatment-resistant depression. *J. Clin. Psychiatry* **68**, 17–25 (2007).
- Burcusa, S. L. & Iacono, W. G. Risk for recurrence in depression. *Clin. Psychol. Rev.* **27**, 959–985 (2007).
- Tranter, R., O'Donovan, C., Chandarana, P. & Kennedy, S. Prevalence and outcome of partial remission in depression. *J. Psychiatry Neurosci* **27**, 241–247 (2002).
- Greden, J. F., Riba, M. B., McInnis, M. G. Treatment resistant depression: a roadmap for effective care. 1st ed. Washington, DC: American Psychiatric Pub.; 2011. xxv, 338pp.
- Frodl, T. et al. Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. *J. Psychiatry Neurosci* **33**, 423–430 (2008).

12. Frodl, T. S. et al. Depression-related variation in brain morphology over 3 years: effects of stress? *Arch. Gen. Psychiatry* **65**, 1156–1165 (2008).
13. Serra-Blasco M, et al. Naturalistic course of major depressive disorder predicted by clinical and structural neuroimaging data: a 5-year follow-up. *Depress. Anxiety* **13**, 1055–1064 (2016).
14. Schmaal, L. et al. Predicting the naturalistic course of major depressive disorder using clinical and multimodal neuroimaging information: a multivariate pattern recognition study. *Biol. Psychiatry* **78**, 278–286 (2015).
15. Farb, N. A., Anderson, A. K., Bloch, R. T. & Segal, Z. V. Mood-linked responses in medial prefrontal cortex predict relapse in patients with recurrent unipolar depression. *Biol. Psychiatry* **70**, 366–372 (2011).
16. Husain, M. M. et al. Speed of response and remission in major depressive disorder with acute electroconvulsive therapy (ECT): a Consortium for Research in ECT (CORE) report. *J. Clin. Psychiatry* **65**, 485–491 (2004).
17. Jelovac, A., Kolshus, E. & McLoughlin, D. M. Relapse following successful electroconvulsive therapy for major depression: a meta-analysis. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol* **38**, 2467–2474 (2013).
18. Hamilton, M. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* **23**, 56–62 (1960).
19. Nierenberg, A. A. & DeCecco, L. M. Definitions of antidepressant treatment response, remission, nonresponse, partial response, and other relevant outcomes: a focus on treatment-resistant depression. *J. Clin. Psychiatry* **62**, 5–9 (2001).
20. Zimmerman, M., Martinez, J. H., Young, D., Chelminski, I. & Dalrymple, K. Severity classification on the hamilton depression rating scale. *J. Affect. Disord.* **150**, 384–388 (2013).
21. Tisdall, M. D. et al. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med./ Soc. Magn. Reson. Med.* **68**, 389–399 (2012).
22. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).
23. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).
24. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
25. Caruana R, Karampatziakis N, Yessinalina A, editors. An empirical evaluation of supervised learning in high dimensions. in *Proceedings of the 25th international conference on Machine learning (ACM)*, 2008.
26. Zhou, X. & Tuck, D. P. MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data. *Bioinformatics* **23**, 1106–1114 (2007).
27. Kuhn, M. Building Predictive Models in R Using the caret Package 28, 26 (2008).
28. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna: Austria, 2014.
29. Brodersen K H, Ong C S, Stephan K E, Buhrmann J M. The balanced accuracy and its posterior distribution. in *Proceedings of the 2010 20th International Conference on Pattern Recognition*. 3121–4.
30. Nadeem M. S. A., Zucker J.-D., Hanczar B, editors. Accuracy-rejection curves (ARCs) for comparing classification methods with a reject option. in *The Proceedings of Machine Learning Research*; 2010.
31. Alfano, K. M. & Cimino, C. R. Alteration of expected hemispheric asymmetries: valence and arousal effects in neuropsychological models of emotion. *Brain Cogn.* **66**, 213–220 (2008).
32. Mneimne, M. et al. Emotional valence and arousal effects on memory and hemispheric asymmetries. *Brain Cogn.* **74**, 10–17 (2010).
33. Herrington, J. D. et al. Localization of asymmetric brain function in emotion and depression. *Psychophysiology* **47**, 442–454 (2010).
34. Koolschijn, P. C., van Haren, N. E., Lensvelt-Mulders, G. J., Hulshoff Pol, H. E. & Kahn, R. S. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Hum. Brain Mapp.* **30**, 3719–3735 (2009).
35. Miller, C. H., Hamilton, J. P., Sacchet, M. D. & Gotlib, I. H. Meta-analysis of functional neuroimaging of major depressive disorder in youth. *JAMA Psychiatry* **72**, 1045–1053 (2015).
36. Downar, J. et al. Anhedonia and reward-circuit connectivity distinguish non-responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biol. Psychiatry* **76**, 176–185 (2014).
37. Lee, W. H. et al. Regional electric field induced by electroconvulsive therapy in a realistic finite element head model: influence of white matter anisotropic conductivity. *NeuroImage* **59**, 2110–2123 (2012).
38. O'Connor, M. K. et al. The influence of age on the response of major depression to electroconvulsive therapy: a C.O.R.E. report. *Am. J. Geriatr. Psychiatry* **9**, 382–390 (2001).
39. Van de Velde, S., Bracke, P. & Levecque, K. Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Soc. Sci. Med.* **71**, 305–313 (2010).
40. Drevets, W. C., Price, J. L. & Furey, M. L. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain. Struct. Funct.* **213**, 93–118 (2008).
41. Wang, T. et al. Disorganized cortical thickness covariance network in major depressive disorder implicated by aberrant hubs in large-scale networks. *Sci Rep.* **6**, 27964 (2016).
42. Vann, S. D., Aggleton, J. P. & Maguire, E. A. What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* **10**, 792–802 (2009).
43. Peng, W., Chen, Z., Yin, L., Jia, Z. & Gong, Q. Essential brain structural alterations in major depressive disorder: a voxel-wise meta-analysis on first episode, medication-naïve patients. *J. Affect. Disord.* **199**, 114–123 (2016).
44. Du, M. et al. Brain grey matter volume alterations in late-life depression. *J. Psychiatry Neurosci.* **39**, 397–406 (2014).
45. Sexton, C. E., Mackay, C. E. & Ebmeier, K. P. A systematic review and meta-analysis of magnetic resonance imaging studies in late-life depression. *Am. J. Geriatr. Psychiatry* **21**, 184–195 (2013).
46. Zhang, K. et al. Molecular, functional, and structural imaging of major depressive disorder. *Neurosci. Bull.* **32**, 273–285 (2016).
47. Wade, B. S. et al. Effect of electroconvulsive therapy on striatal morphometry in major depressive disorder. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **41**, 2481–2491 (2016).
48. Schmaal, L. et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol. Psychiatry* **21**, 806–812 (2016).
49. Lorenzetti, V., Allen, N. B., Fornito, A. & Yucel, M. Structural brain abnormalities in major depressive disorder: A selective review of recent MRI studies. *J. Affect. Disord.* **117**, 1–17 (2009).
50. Disner, S. G., Beevers, C. G., Haigh, E. A. & Beck, A. T. Neural mechanisms of the cognitive model of depression. *Nat. Rev. Neurosci.* **12**, 467–477 (2011).
51. Ochsner, K. N., Silvers, J. A. & Buhle, J. T. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann. N. Y. Acad. Sci.* **1251**, E1–E24 (2012).
52. Kuhn, S., Vanderhasselt, M. A., De Raedt, R. & Gallinat, J. The neural basis of unwanted thoughts during resting state. *Soc. Cogn. Affect. Neurosci.* **9**, 1320–1324 (2014).
53. Sowell, E. R. et al. Mapping cortical change across the human life span. *Nat. Neurosci.* **6**, 309–315 (2003).
54. Wong, D. F. et al. Effects of age on dopamine and serotonin receptors measured by positron tomography in the living human brain. *Science* **226**, 1393–1396 (1984).
55. Sexton, C. E. et al. Magnetic resonance imaging in late-life depression: vascular and glucocorticoid cascade hypotheses. *Br. J. Psychiatry: J. Mental Sci.* **201**, 46–51 (2012).
56. Taylor, W. D., Aizenstein, H. J. & Alexopoulos, G. S. The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Mol. Psychiatry* **18**, 963–974 (2013).
57. Fradkin Y, Khadka S, Bessette K L, Stevens M C. The relationship of impulsivity and cortical thickness in depressed and non-depressed adolescents. *Brain. Imaging Behav* **11**, 1515–1525 (2016).
58. Radua, J. et al. Neural response to specific components of fearful faces in healthy and schizophrenic adults. *NeuroImage* **49**, 939–946 (2010).
59. Oathes, D. J. & Ray, W. J. Depressed mood, index finger force and motor cortex stimulation: a transcranial magnetic stimulation (TMS) study. *Biol. Psychol.* **72**, 271–277 (2006).
60. Maeda, F., Keenan, J. P. & Pascual-Leone, A. Interhemispheric asymmetry of motor cortical excitability in major depression as measured by transcranial magnetic stimulation. *Br. J. Psychiatry: J. Mental Sci.* **177**, 169–173 (2000).
61. Chistyakov, A. V. et al. Effect of electroconvulsive therapy on cortical excitability in patients with major depression: a transcranial magnetic stimulation study. *Clin. Neurophysiol.* **116**, 386–392 (2005).

62. Bhagwagar, Z. et al. Reduction in occipital cortex gamma-aminobutyric acid concentrations in medication-free recovered unipolar depressed and bipolar subjects. *Biol. Psychiatry* **61**, 806–812 (2007).
63. Grieve, S. M., Korgaonkar, M. S., Koslow, S. H., Gordon, E. & Williams, L. M. Widespread reductions in gray matter volume in depression. *NeuroImage Clin.* **3**, 332–339 (2013).
64. Liao, Y. et al. Is depression a disconnection syndrome? Meta-analysis of diffusion tensor imaging studies in patients with MDD. *J Psychiatry Neurosci.* **38**, 49–56 (2013).
65. Meng, C. et al. Aberrant topology of striatum's connectivity is associated with the number of episodes in depression. *Brain* **137**, 598–609 (2014).
66. Na, K. S. et al. Brain-derived neurotrophic factor promoter methylation and cortical thickness in recurrent major depressive disorder. *Sci. Rep.* **6**, 21089 (2016).
67. Sanacora, G. et al. Cortical gamma-aminobutyric acid concentrations in depressed patients receiving cognitive behavioral therapy. *Biol. Psychiatry* **59**, 284–286 (2006).
68. Sanacora, G. et al. Increased cortical GABA concentrations in depressed patients receiving ECT. *Am. J. Psychiatry* **160**, 577–579 (2003).
69. Sanacora, G., Mason, G. F., Rothman, D. L. & Krystal, J. H. Increased occipital cortex GABA concentrations in depressed patients after therapy with selective serotonin reuptake inhibitors. *Am. J. Psychiatry* **159**, 663–665 (2002).
70. Bouckaert, F. et al. ECT: its brain enabling effects: a review of electroconvulsive therapy-induced structural brain plasticity. *J. ECT* **30**, 143–151 (2014).
71. Joshi, S. H. et al. Structural plasticity of the hippocampus and amygdala induced by electroconvulsive therapy in major depression. *Biol. Psychiatry* **79**, 282–292 (2016).
72. Scott, A. I. Does ECT alter brain structure? *Am. J. Psychiatry* **152**, 1403 (1995).
73. Bolwig, T. G., Hertz, M. M., Paulson, O. B., Spotoft, H. & Rafaelsen, O. J. The permeability of the blood-brain barrier during electrically induced seizures in man. *Eur. J. Clin. Invest.* **7**, 87–93 (1977).