Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data

Ronny Redlich, PhD; Nils Opel, MD; Dominik Grotegerd, MSc; Katharina Dohm, MSc; Dario Zaremba, MSc; Christian Bürger, MSc; Sandra Münker; Lisa Mühlmann, MD; Patricia Wahl, MSc; Walter Heindel, MD; Volker Arolt, MD, PhD; Judith Alferink, MD; Peter Zwanzger, MD, PhD; Maxim Zavorotnyy, MD; Harald Kugel, PhD; Udo Dannlowski, MD, PhD

IMPORTANCE  Electroconvulsive therapy (ECT) is one of the most effective treatments for severe depression. However, biomarkers that accurately predict a response to ECT remain unidentified.

OBJECTIVE  To investigate whether certain factors identified by structural magnetic resonance imaging (MRI) techniques are able to predict ECT response.

DESIGN, SETTING, AND PARTICIPANTS  In this nonrandomized prospective study, gray matter structure was assessed twice at approximately 6 weeks apart using 3-T MRI and voxel-based morphometry. Patients were recruited through the inpatient service of the Department of Psychiatry, University of Muenster, from March 11, 2010, to March 27, 2015. Two patient groups with acute major depressive disorder were included. One group received an ECT series in addition to antidepressants (n = 24); a comparison sample was treated solely with antidepressants (n = 23). Both groups were compared with a sample of healthy control participants (n = 21).

MAIN OUTCOMES AND MEASURES  Binary pattern classification was used to predict ECT response by structural MRI that was performed before treatment. In addition, univariate analysis was conducted to predict reduction of the Hamilton Depression Rating Scale score by pretreatment gray matter volumes and to investigate ECT-related structural changes.

RESULTS  One participant in the ECT sample was excluded from the analysis, leaving 67 participants (27 men and 40 women; mean [SD] age, 43.7 [10.6] years). The binary pattern classification yielded a successful prediction of ECT response, with accuracy rates of 78.3% (18 of 23 patients in the ECT sample) and sensitivity rates of 100% (13 of 13 who responded to ECT). Furthermore, a support vector regression yielded a significant prediction of relative reduction in the Hamilton Depression Rating Scale score. The principal findings of the univariate model indicated a positive association between pretreatment subgenual cingulate volume and individual ECT response (Montreal Neurological Institute [MNI] coordinates x = 8, y = 21, z = −18; Z score, 4.00; P < .001; peak voxel r = 0.73). Furthermore, the analysis of treatment effects revealed an increase in hippocampal volume in the ECT sample (MNI coordinates x = −28, y = −9, z = −18; Z score, 7.81; P < .001) that was missing in the medication-only sample.

CONCLUSIONS AND RELEVANCE  A relatively small degree of structural impairment in the subgenual cingulate cortex before therapy seems to be associated with successful treatment with ECT. In the future, neuroimaging techniques could prove to be promising tools for predicting the individual therapeutic effectiveness of ECT.
Prediction of electroconvulsive therapy (ECT) response has been a scientific goal since the early 1950s. A large number of studies were devoted to ECT response prediction and have suggested indexes, rating scales, and symptom clusters that were proposed to be predictive of treatment success.1-4 However, a systematic approach to develop clinically useful predictors has mostly failed.5,6

Recently, structural neuroimaging techniques have come into focus to provide insights into the underlying neuronal mechanisms of depression and antidepressant treatment effects. Major depressive disorder (MDD) has been repeatedly shown to be associated with structural abnormalities, with volume reductions in the hippocampus and anterior cingulate gyrus among the most consistent findings.7,8

The pertinent questions are whether and to what extent these alterations can be reversed by ECT. In animals, the application of electroconvulsive shock has led to hippocampal neuroplasticity effects,9 probably by inducing neurogenesis,10 gliogenesis,11 endothelial cell proliferation,12 and angiogenesis.13 Increasing evidence suggests that ECT also induces neuroplasticity in the human brain, specifically within the hippocampus and amygdala.14-18 However, whether these neuroplastic effects are associated with clinical response remains unclear. Although one recent study14 reported a positive association between an increase in hippocampal gray matter volume (GMV) and symptomatic improvement,2 other studies15,16 reported no relationship, inverse or otherwise. However, assessing and comparing these heterogeneous findings is difficult because studies vary dramatically in terms of sample size, study design (eg, inclusion15 or exclusion14,16,17,19 of an MDD control group), and magnetic resonance imaging (MRI) analysis (eg, subcortical segmentation14,15 vs voxel-based morphometry15).

Aside from the question of structural changes in the brain induced by ECT, the prediction of treatment response is of high clinical relevance because approximately one-third of patients do not benefit from ECT. A few neuroimaging studies within the past 2 years have addressed the issue of finding biomarkers for ECT response and report that resting-state connectivity,16 larger pretreatment amygdala volume,19 and smaller inferior frontal gyrus volume20 might be associated with more effective ECT. In general, studies have shown that a reduced hippocampal volume might be a risk factor for the development of depression21,22 and is related to a poorer longitudinal clinical outcome in patients with MDD.23

However, because classic univariate group statistics are unable to predict individual cases, multivariate approaches, such as pattern classification techniques, have been suggested as promising tools for overcoming these methodologic problems; successful applications include differentiation between MDD and bipolar disorders24 or the prediction of remission in ECT-treated patients using functional MRI.17 Accurate prediction of ECT response using structural MRI before treatment would be of enormous clinical valuable because structural MRI data are often obtained routinely in clinical practice.

Therefore, this naturalistic study, primarily aligned to investigate ECT-related changes, has been designed to investigate the following objectives:

1. To predict individual ECT response (binary classification and linear prediction) by multivariate pattern classification techniques on structural MRI data (main study objective); and
2. To identify regional brain morphometric biomarkers that are associated with the degree of symptom relief in patients treated with ECT (secondary study objective 1) and to detect longitudinal effects of ECT on brain structure using univariate statistics (secondary study objective 2).

### Methods

#### Participants and Study Design

The present nonrandomized prospective study included 47 patients with acute MDD (Hamilton Depression Rating Scale25 [HDRS] score range, 14 to 42), divided according to treatment into a sample undergoing ECT (n = 24) and a sample receiving medication only (n = 23). Patients were recruited from March 11, 2010, to March 27, 2015. A healthy control sample (n = 21) was additionally recruited and matched to both patient groups by age, sex, and educational level (Table 1). One patient in the ECT sample had to be excluded owing to anatomical abnormalities. This study was approved by the institutional review board of the University of Muenster, and all participants provided written informed consent before participation.

All patients underwent scanning twice, with each session occurring a mean (SD) of 6.3 (1.6) weeks apart (Table 1). Patients were recruited through the inpatient service of the Department of Psychiatry, University of Muenster. The treatment form was based on clinical decisions independent from study participation. Patients in both groups received antidepressants at both time points, and this therapy was not discontinued in the ECT sample. None of the patients received benzodiazepines. Diagnoses were verified using the German version of the Structured Clinical Interview for DSM-IV.26 All patients had a current major depressive episode and fulfilled the criteria for MDD. Questionnaire measures and medication load are described in eMethods 1 in the Supplement. For the healthy control sample, any lifetime psychiatric disorder was an exclusion criterion. Further exclusion criteria for all participants were any neurologic abnormalities, organic mental disorders, dementia, brain injuries, or contraindications to MRI.

### Key Points

**Question** Can structural magnetic resonance imaging (MRI) techniques be used to predict electroconvulsive therapy (ECT) response?

**Findings** In this nonrandomized prospective study of patients with severe depression, we achieved a successful prediction of ECT response, with accuracy rates as great as 78.3%, using structural MRI obtained before therapy. Furthermore, ECT induced a massive increase in hippocampal volume, and a relatively small degree of structural impairment in the subgenual cingulate cortex before therapy seemed to be associated with successful ECT.

**Meaning** Neuroimaging techniques may be useful for predicting the individual therapeutic effectiveness of ECT.
For patients with MDD, comorbid lifetime substance-related disorders, bipolar disorders, schizophrenia, and other psychotic disorders were exclusion criteria (eTable in the Supplement). No significant differences in frequencies of comorbid diagnoses were found between the ECT and medication-only samples ($P > .05$).

**Electroconvulsive Therapy**

Brief-pulse ECT was conducted 3 times a week using an integrated instrument (Thymatron system IV; Somatics Inc). Initially, 9 to 12 sessions of ECT were given, and sessions were continued if the patients did not achieve symptom relief (mean [SD] No. of sessions, 14.0 [3.8]; range, 9–24). A detailed description of ECT treatment can be found in eMethods 2 in the Supplement.

### Data Acquisition

The data sets were acquired using a 3-T scanner and previously published protocols. A detailed description of data collection and voxel-based morphometry can be found in eMethods 3 in the Supplement.

### Pattern Classification

Pattern classification approaches include a set of machine learning–based algorithms that allow multivariate differentiation of 2 or more groups based on high-dimensional data, such as structural brain images. Although the univariate approach is based on independent voxelwise processing of data, the multivariate approach is based on the recognition of the patterns of the voxels. Two well-established and frequently used classifiers in neuroimaging were applied on smoothed,
modulated, whole-brain gray matter data without masking or other feature selection. First, a support vector machine (SVM) was applied, which has been used successfully in differentiating patients with depression from healthy controls and patients with unipolar disorder from those with bipolar disorder using neuroimaging data. A linear kernel was used with default settings and the parameter $C = 1$, which is recommended for high-dimensional data and relatively small sample sizes. Second, a Gaussian process classifier (GPC) was applied to validate results with an independent algorithm. Pattern classification analyses were conducted using the MANIA toolbox with the LIBSVM liblinear implementation and the GPC.

To examine our main objective, the ECT sample was classified as responders ($n = 13$) and nonresponders ($n = 10$), with the criterion of less than 50% individual symptom relief for nonresponse according to HDRS scores. Responders and nonresponders did not differ with regard to acute symptom severity as measured by the Beck Depression Inventory and HDRS scores ($P = .71$) or any other clinical parameter from Table 1 ($P > .11$). We evaluated the ability of the SVM and GPC to predict therapy response using pretreatment gray matter data with a leave-one-subject-out cross-validation procedure. The corresponding weight vectors of the SVM hyperplane were extracted to visualize contributing brain areas. The test margin, denoting the confidence with which individual participants were classified as responders or nonresponders, was extracted from each patient to investigate the influence of medication load on individual-level classification performance. Statistical significance was empirically estimated by performing permutation tests (1000 iterations) as suggested by the literature. In addition, we performed the same classification procedures within the medication-only sample.

To predict a continuous improvement in symptoms measured by the individual symptom relief according to the HDRS within the entire ECT sample, a support vector regression (SVR) was conducted based on an anatomical mask of the subgenual cingulate gyrus according to Talairach atlas definitions. Therefore, an epsilon SVR, based on the LIBSVM implementation with linear kernel and default parameter, was applied using a leave-one-subject-out cross-validation procedure.

### Univariate Analysis

Group statistics were calculated using statistical parameter mapping (SPM8; Welcome Department of Cognitive Neurology). To investigate our secondary research objectives, we pursued 2 analysis strategies. First, to detect brain area volumes that are associated with symptom relief (percentage of change in individual HDRS scores) in patients treated with ECT on the group level, a regression model was created regressing the pretreatment whole-brain GMVs on the individual percentage of HDRS changes ($1 - \frac{\text{HDRS at Time 2}}{\text{HDRS at Time 1}}$). We maintained a corrected false-positive detection rate of $P < .05$ using a voxel threshold of $P < .001$ and an empirically determined cluster extent threshold ($k$ statistic) determined by Monte Carlo simulations (1000 iterations) yielding $k = 403$ voxels. Automated anatomical labeling was performed by means of the AAL toolbox.

Second, to investigate structural variation related to diagnosis and treatment effects, we created a full factorial model that included the between-participants factor group (ECT vs medication-only vs healthy control samples) and the within-participants factor time (times 1 vs 2). Potential interactions were followed by appropriate unpaired $t$ tests for comparing samples and paired $t$ tests for comparing time points. All analyses were corrected for multiple comparisons with the same parameter described above ($k = 403$ voxels). Other analyses include the association between gray matter changes (time 2 – time 1) and percentage of HDRS improvement and associations between symptom improvement and clinical and sociodemographic data (eMethods 5 in the Supplement).

### Table 2. Prediction of ECT Response by GMV Obtained Before Therapy

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy, %</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>No. Identified/No. of Participants</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIBSVM</td>
<td>78.3</td>
<td>100</td>
<td>50.0</td>
<td>13/13 5/10</td>
<td>.009</td>
</tr>
<tr>
<td>GPC</td>
<td>73.9</td>
<td>100</td>
<td>40.0</td>
<td>13/13 4/10</td>
<td>.01</td>
</tr>
</tbody>
</table>

Abbreviations: ECT, electroconvulsive therapy; GMV, gray matter volume; GPC, Gaussian process classifier; LIBSVM, library for support vector machines.

### Results

#### Prediction of ECT Response by Means of Pattern Classification

Sixty-seven participants (27 men; 40 women; mean [SD] age, 43.7 [10.6] years) were included in the analysis. The binary and linear prediction of ECT response by structural images obtained before treatment (the primary objective) yielded accuracy rates significantly above the level of chance (Table 2). The SVM was able to differentiate ECT responders and nonresponders with an accuracy rate of 78.3% (18 of 23 in the ECT sample; $P = .009$; sensitivity, 100%; [13 of 13 responders]; specificity, 50.0% [5 of 10 nonresponders]). When using the GPC, a similar accuracy rate was obtained at 73.9% (17 of 23 in the ECT sample; $P = .04$; sensitivity, 100%; [13 of 13 responders]; specificity, 40.0% [4 of 10 nonresponders]). The discriminative maps of the dichotomous classifications show the subgenual cingulate gyrus to be the area that contributed most to the classification of therapy response (eFigure in the Supplement). No sig-
significant associations were found between medication load, age, or sex and pattern classification test marginals (P ≥ .68). The SVR results were significantly associated with continuous symptom relief according to the HDRS using the whole ECT sample (r = 0.67; P < .001). The mean (SD) difference between the predicted change in HDRS score and the actual percentage of change in HDRS score was 19.2% (13.1%). A detailed illustration of the results of the SVR can be found in Figure 1.

Association of ECT Response With Whole-Brain Gray Matter Data
The whole-brain regression analysis (secondary study objective 1) yielded only 1 significant cluster, showing a strong positive association of subgenual cingulate gyrus volume at time 1 and ECT response (Montreal Neurological Institute [MNI] coordinates x = 8, y = 21, z = −18; Z = 4.00; P < .001; peak voxel r = 0.73; cluster size, k = 873 voxels). In other words, higher pretreatment subgenual cingulate gyrus GMV is associated with a better clinical response (Figure 2). This finding corresponded well with our multivariate results (eFigure in the Supplement), which showed the subgenual cingulate gyrus to be the area with the most significant contribution to the classification of therapy response. No significant associations between GMV at time 1 and symptom relief were found in the medication-only sample.

Full Factorial Model Group × Time Interaction
Cross-sectional Effects of Group at Time 1
In both patient samples, reduced GMVs at time 1 were found in neural regions previously reported as showing abnormalities in individuals with unipolar depression.7,8,45 These areas included the anterior cingulate gyrus, hippocampus, temporal areas, and precuneus compared with the control group (details can be found in eResults 1 in the Supplement).

Longitudinal Effects of Treatment on GMVs
Both patient samples responded significantly, as measured by the mean (SD) change in the HDRS score at time 1 (ECT sample, 26.0 [6.5]; medication-only sample, 23.1 [4.8]) and at time 2 (ECT sample, 13.1 [7.1]; medication-only sample, 12.9 [7.4]; P < .001 for differences between times for both samples). Improvements in symptom severity in both patient groups did not significantly correlate with age (P ≥ .39), sex (P ≥ .11), HDRS scores (P ≥ .35), illness characteristics (P ≥ .11), or medication load index (P ≥ .12).

The analysis of the GM data using the 3 × 2 full-factorial model (secondary study objective 2) yielded a strong interaction effect of group × time within the left hippocampus (MNI coordinates x = −33, y = −7, z = −24; F(2,64) = 30.71; Z = 6.67; k = 5660 voxels; P < .001), resulting from an increase in gray matter voxels in the ECT sample over time (MNI coordinates x = −28, y = −9, z = −18; Z score, 7.81; P < .001), whereas the GMVs of the medication-only and healthy control samples did not change significantly (Figure 3). The post hoc analyses of cross-sectional effects at time 2 revealed that whole-brain GMV reductions in the ECT sample at time 1 were entirely normalized after ECT treatment, compared with the healthy control sample. Moreover, the ECT sample showed a slightly increased hippocampal volume and GMV at time 2 compared with the healthy control sample. Details and additional post hoc tests can be found in eResults 2 in the Supplement.

Discussion
The present study sought to predict ECT response in a psychiatric sample by using a combination of structural MRI data and machine learning techniques. The prediction of ECT response was successful with high rates of sensitivity. The principal findings of the regression analysis indicated positive associations between pretreatment subgenual cingulate volume and ECT response. Furthermore, the analyses of longitudinal data revealed an increase in bilateral hippocampal volume solely in the ECT sample, whereas evidence for such neuroplastic effects was absent in the medication-only sample.

The ability to advise psychiatrists and patients accurately regarding the chances of successful ECT is of considerable value, particularly because ECT is a demanding procedure and, despite having relatively few adverse effects, has a profound effect on patients. The binary pattern classification analyses yielded accuracy rates as great as 78%. The sensitivity rate is particularly remarkable because every responder was correctly classified as such. Naturally, this sensitivity was accompanied by lower specificity rates of 50.0% and 40.0%, implying that 5 of 10 and 4 of 10 nonresponders were classified as responders. This finding is probably partially owing to 6 patients who were designated as nonresponders although they had partial responses (HDRS score changes in the range of 27%-46%). Support vector regression might be a suitable approach to overcome this issue because it provides a continuous prediction of symptom improvement. However, a linear...
classification output requires a decision with regard to whether a predicted symptom improvement of approximately 30% to 50% justifies ECT or not.

Our univariate whole-brain regression analyses revealed that, in particular, 1 area in the brain seems to have a decisive effect on the prediction of clinical response to ECT: the GMV of the subgenual cingulate gyrus before treatment was positively associated with ECT response. Moreover, corresponding to this finding, the discriminative maps of the binary multivariate classification show the subgenual cingulate gyrus to be the area with the most significant contribution to the classification of ECT response. The subgenual cingulate gyrus is known to be abnormal in function and structure in patients with MDD and is believed to have a key role in the processing and regulation of emotions. The subgenual cingulate gyrus is also one of the most effective targets in deep brain stimulation studies of treatment-resistant depression. In our study, a decrease in the volume of the subgenual cingulate gyrus before treatment indicates a poorer clinical outcome after ECT. One possible explanation could be that, instead of normalizing the gray matter structure, successful treatment might result in a relative normalization of the connectivity pattern of the subgenual cingulate gyrus to other limbic areas, which in turn might require a less impaired subgenual cingulate structure. However, because these interpretations are speculative, future studies should test this hypothesis by using functional connectivity MRI, for example.

Figure 2. Association of Pretreatment Subgenual Anterior Cingulate Gray Matter Volume (GMV) and Symptom Relief

A. Sagittal view depicts the positive association between the percentage of change in the Hamilton Depression Rating Scale (HDRS) score and subgenual anterior cingulate volume before electroconvulsive therapy (ECT) (peak voxel r = 0.73). Statistics were corrected for the entire brain volume (P < .001, k = 403), yielding only 1 significant cluster mapping to this area. B. Scatterplot depicts the positive correlation of the subgenual cingulate gyrus GMVs and the percentage of symptom improvement as measured by HDRS scores (ECT response). Contour lines indicate individual 95% CIs; central diagonal line indicates the association between subgenual anterior cingulate volume before ECT and percentage of change in the HDRS score.

Figure 3. Longitudinal Effects of Electroconvulsive Therapy (ECT) on Whole-Brain Gray Matter Volume (GMV)

The sagittal and coronal sections (coordinates according to Montreal Neurological Institute space) feature the GMV increases in the ECT group mapping predominantly to the hippocampal formation. A corrected false-positive detection rate of P < .05 using a voxel threshold of P < .001 with an empirically determined cluster extent threshold (k), determined by Monte Carlo simulations (yielding k = 403 voxels), was used.
Our analysis of the longitudinal effects of ECT confirmed previous findings\textsuperscript{14-16,18} that ECT induces massive structural plasticity, particularly in the hippocampus. Synaptogenesis, dendrogenesis, angiogenesis, or neurogenesis, as already shown in animal models,\textsuperscript{10-13} have been suggested as processes that mediate the observed effects.\textsuperscript{14} Furthermore, the results might reflect changes in blood flow or volume, although the anatomical reasons for local GMV changes remain unclear. However, in the present study, the extent of this hippocampal plasticity did not positively correlate to the extent of clinical response, as already reported,\textsuperscript{15,16} which indicates a possibility that these GMV increases are a byproduct rather than the underlying mechanism of ECT.\textsuperscript{31} On the other side, the missing link between GMV increases and symptom improvement may underlie long-term neurostructural effects that are associated not with acute but with latent symptomatic improvement. To date, whether structural changes can explain the therapeutic efficacy remains unclear, and the findings reported by recent studies are very inconsistent and often contradictory. The need for replication studies is urgent, particularly with regard to imaging measures and mood responses.

Despite these first encouraging results, some limitations must be acknowledged. First, patients were not randomized to treatment groups owing to the naturalistic study design. Although both treatment groups shared several similarities, they differed regarding medication and some illness characteristics and therefore were not comparable. Furthermore, the analysis of the longitudinal effects of medication in the non-ECT sample did not show any significant volume changes during treatment, and our pattern classification approach could not predict therapy outcome significantly. On one hand, ECT may induce more pronounced changes in structure, unlike antidepressants, which may induce only subtle changes. Methodologic reasons might include a heterogeneous medical treatment of the medication-only sample, whereas the ECT sample was uniquely treated with respect to ECT. Furthermore, the medication-only sample had already received medication before the study. Therefore, no meaningful conclusions can be drawn with regard to the treatment effects in this clinical control group per se. In addition, all patients in the ECT group received medication that might have influenced the reported effects. On the other hand, because the changes in volume were completely lacking in the medication-only group, these effects were presumably induced by ECT; furthermore, medication load was not associated with classification performance. This finding is supported by a recent study showing that changes in volume occur in patients withdrawing from psychotropic medication therapy before ECT and thus are independent of pharmacotherapy.\textsuperscript{14} Nevertheless, pharmacologically induced changes might still have contributed to the reported results. Finally, our findings are strongly in need of replication in an independent validation sample. Besides replication, future studies should also aim to transfer a trained classifier to independent sites.

Conclusions

The present study might be a promising step to provide biomarkers for correctly identifying patients who are likely to respond to ECT. Although determining which ECT recipients will respond remains difficult in clinical practice, a routine assessment with structural MRI before treatment could serve as a decision guide for clinical psychologists.
Research

Original Investigation

Prediction of Electroconvulsive Therapy Response in Severe Depression


