Predicting Treatment Response to Cognitive Behavioral Therapy in Panic Disorder With Agoraphobia by Integrating Local Neural Information

Tim Hahn, PhD; Tilo Kircher, MD; Benjamin Straube, PhD; Hans-Ulrich Wittchen, PhD; Carsten Konrad, MD; Andreas Ströhle, MD; André Wittmann, PhD; Bettina Pfleiderer, MD, PhD; Andreas Reif, MD; Volker Arolt, MD; Ulrike Lueken, PhD

IMPORTANCE Although neuroimaging research has made substantial progress in identifying the large-scale neural substrate of anxiety disorders, its value for clinical application lags behind expectations. Machine-learning approaches have predictive potential for individual-patient prognostic purposes and might thus aid translational efforts in psychiatric research.

OBJECTIVE To predict treatment response to cognitive behavioral therapy (CBT) on an individual-patient level based on functional magnetic resonance imaging data in patients with panic disorder with agoraphobia (PD/AG).

DESIGN, SETTING, AND PARTICIPANTS We included 49 patients free of medication for at least 4 weeks and with a primary diagnosis of PD/AG in a longitudinal study performed at 8 clinical research institutes and outpatient centers across Germany. The functional magnetic resonance imaging study was conducted between July 2007 and March 2010.

INTERVENTIONS Twelve CBT sessions conducted 2 times a week focusing on behavioral exposure.

MAIN OUTCOMES AND MEASURES Treatment response was defined as exceeding a 50% reduction in Hamilton Anxiety Rating Scale scores. Blood oxygenation level–dependent signal was measured during a differential fear-conditioning task. Regional and whole-brain gaussian process classifiers using a nested leave-one-out cross-validation were used to predict the treatment response from data acquired before CBT.

RESULTS Although no single brain region was predictive of treatment response, integrating regional classifiers based on data from the acquisition and the extinction phases of the fear-conditioning task for the whole brain yielded good predictive performance (accuracy, 82%; sensitivity, 92%; specificity, 72%; P < .001). Data from the acquisition phase enabled 73% correct individual-patient classifications (sensitivity, 80%; specificity, 67%; P < .001), whereas data from the extinction phase led to an accuracy of 74% (sensitivity, 64%; specificity, 83%; P < .001). Conservative reanalyses under consideration of potential confounders yielded nominally lower but comparable accuracy rates (acquisition phase, 70%; extinction phase, 71%; combined, 79%).

CONCLUSIONS AND RELEVANCE Predicting treatment response to CBT based on functional neuroimaging data in PD/AG is possible with high accuracy on an individual-patient level. This novel machine-learning approach brings personalized medicine within reach, directly supporting clinical decisions for the selection of treatment options, thus helping to improve response rates.
Although neuroimaging research has made substantial progress in identifying the neural substrate of anxiety disorders, its value for clinical translational research lags behind expectations. One reason is that most studies are based on conventional group analyses. In contrast, a useful biomarker has to provide sufficient sensitivity and specificity to predict a given patient’s status on the individual level. Machine-learning algorithms, such as Support Vector Machines or gaussian process classifiers (GPC), have shown predictive potential for single-subject diagnostic purposes.

Although results of such proof-of-concept studies appear highly encouraging, predictive approaches need to facilitate clinical decisions. To date, however, most studies that use machine learning merely predict a diagnostic label (a review and notable exceptions appear in Klöppel et al9 and Sundermann et al10). Among such relevant clinical decisions, the prediction of treatment outcome is of utmost importance because identifying individuals who will not benefit from a therapeutic intervention enables early treatment modification and may offer better outcomes for otherwise nonresponsive patients.

Evidence from the field of mood disorders suggests that neuroimaging biomarkers may aid in classifying disease status and predicting response to pharmacological treatment or cognitive behavioral therapy (CBT). However, despite the large number of neuroimaging studies on anxiety disorders, analyses on response markers in this patient group are surprisingly scarce. Application of regression methods to neuroimaging data from social anxiety disorder has resulted in improved rates of explained variance compared with clinical outcome parameters. These improved rates have been reported when adding neuroimaging data (12% vs 41%) that were tested in a cross-validation procedure using machine-learning algorithms.

As one of the most disabling anxiety disorders, panic disorder (PD) is associated with high individual and societal burdens. Although CBT has been proven efficient and is recommended as a first-line treatment for PD, response rates are far from satisfying. Recent data from the national research initiative PANIC-NET originating from the multicenter trial Mechanism of Action in CBT on patients with PD with agoraphobia (PD/AG) who were treated with exposure-based CBT indicated positive response rates in approximately half of the sample. Comparable effect sizes have been reported for other randomized clinical trials as shown in a meta-analysis by Sánchez-Meca et al. Identifying patients with a higher risk for nonresponse before treatment could allow for individualized treatment decisions and thus improve response rates.

Fear conditioning is a core process for the development and maintenance of PD/AG. The nature of the learning deficit remains poorly described, but studies point toward the relevance of discriminatory learning, overgeneralization of fear, and attenuated extinction learning. On a neural level, brain systems mediating threat in response to stimuli that signal safety have been reported as a pathophysiological correlate of PD/AG. In line with this correlate, results in a subset of PD/AG patients from the previous study indicate that, before treatment, nonresponse was characterized by enhanced safety-signal processing in brain systems associated with the processing of threat (e.g., pregenual anterior cingulate cortex, right hippocampus, and right amygdala) and altered anterior cingulate cortex-amygdala coupling.

Although these findings are informative about phenotype characteristics related to treatment response, they are not applicable to the individual patient. Herein we aim to predict individual responses to CBT in the previously described PD/AG sample, thereby providing information of high clinical relevance based on functional magnetic resonance imaging (fMRI) data. We developed a novel multivariate pattern classification approach with regional patterns of neural responses while allowing for the integration of local predictions in the whole brain, thus considering local and whole-brain neural information to facilitate the prediction of treatment response in individual patients.

Methods

Multicenter Mechanism of Action in CBT Study

This work is part of the German multicenter trial Mechanism of Action in CBT. After a complete description of the study protocol, written informed consent was obtained from the participants, and the protocol was approved by the local ethics committees in each fMRI center according to the Declaration of Helsinki. The randomized clinical trial (isrctn.org identifier: ISRCTN80046034) was approved by the ethics committee of the Medical Faculty of the Technische Universität Dresden (agreement EK 164082006). The neuroimaging components were approved by the ethics committee of the Medical Faculty of the RWTH Aachen University, Aachen (agreement EK 073/07) and at all local sites. The experimental pharmacology study was approved by the ethics committee of the state of Berlin (EudraCT 2006-00-4860-29).

Within the German psychotherapy research network PANIC-NET, a multicenter randomized clinical trial of CBT for PD/AG patients was conducted in 8 centers across Germany. Three hundred sixty-nine patients free of medication for at least 4 weeks who met DSM-IV-TR criteria for PD/AG were treated with a manualized treatment protocol or assigned to a waiting list that consisted of 12 sessions of CBT 2 times a week focusing on behavioral exposure in situ. Two procedural variants of CBT (identical in content and dose) were compared that differed only with regard to therapist guidance during exposure sessions. In the therapist-guided condition, the therapist accompanied the patient during exposure, whereas patients were instructed by the therapist but performed the exposure on their own in the nonguided condition. Patients were randomly assigned to 1 of the 2 CBT arms; both groups exhibited significant symptom reduction after CBT; therefore, they were combined in the fMRI responder analyses (Table 1). Response was defined as a reduction in Hamilton Anxiety Rating Scale scores exceeding 50% from baseline to posttreatment assessment. Details on the study design and immediate and long-term treatment effects have been reported previously (Appendix in the Supplement). Four of the 8 centers participated in an fMRI add-on study, yielding 49 quality-controlled fMRI data sets at baseline. The fMRI study was conducted between July 2007 and March 2010. Patient
The task, fMRI data acquisition, and preprocessing pathways have been described in detail. We applied a differential fear-conditioning task. Colored geometric stimuli served as conditioned stimuli (CS), meaning that the white noise, 100 milliseconds) was used as the unconditioned stimulus (US), which was pseudorandomly paired with a CS during the acquisition phase, whereas the unpaired CS was never followed by the US (counterbalanced among patients; reinforcement rate, 50%). During acquisition, only those trials in which no US was delivered (CS+) were analyzed. Image acquisition and analysis pathways are described in the eAppendix in the Supplement. For the GPC analyses, we used contrast maps from individual patients to calculate the overall prediction accuracy of this approach, a nested comparisons (for 55 regions) using the false discovery rate.

**Gaussian Process Classification**

**Regional GPCs**

Using the contrast images for the acquisition and the extinction phases, whole-brain data from 55 regions drawn from the Harvard-Oxford Brain Atlas as described in Carter et al were extracted for each patient. These data were analyzed as described by Marquand et al using GPCs (details are given in the eAppendix in the Supplement). We predicted a patient’s probability to be a responder independently for each region based on all voxels within the respective region using leave-one-out cross-validation.

We evaluated the performance of the 55 regional classifiers by converting the predictive probabilities to categorical predictions. We applied a threshold that categorized a patient as a responder or a nonresponder if his or her probability of response was larger than 0.5 or smaller than 0.5, respectively. Accuracies were calculated as the ratio of correct predictions to the number of patients for each GPC. Because the numbers of responders and nonresponders were not the same, classification accuracies were calibrated. We established whether these regional accuracies were statistically significant, we ran each classifier 1000 times with randomly permuted labels and counted the number of permutations that achieved higher accuracy than the one observed with the true labels. We calculated the P value by dividing this number by 1000. We corrected for multiple comparisons (for 55 regions) using the false discovery rate.

**Integration of Regional Predictive Probabilities**

We applied a second classifier with the same specifications used for the local classifiers to integrate the predictive classification probabilities obtained from the local leave-one-out classifiers. Those probabilities were used as predictors for the algorithm in order to determine the classification of a participant. To calculate the overall prediction accuracy of this approach, a nested leave-one-out procedure was implemented (details are given in the eAppendix in the Supplement). This procedure ensures complete independence of the training and the test data set.

The significance of the whole-brain classifier’s prediction accuracy was tested in analogy to the regional classifiers. To compare the best local classifier with the whole-brain classifier, we used a permutation-based procedure. We first calculated the difference between the best local classifier and the whole-brain classifier. Then, we approximated the null distribution of this difference by computing the difference between the 2 respective classifiers based on 1000 accuracies under permutation for the best local classifier and the whole-brain classifier. Finally, we deemed 2 accuracies significantly different if less than 5% of the absolute accuracy differences
under permutation were larger than the difference obtained with the true labels.

Regional and Multivariate Mapping
To determine those brain regions that contributed most to classification, we derived weight maps from the GPC models as described in Marquand et al.36 and calculated the means for all cross-validation folds. Although this procedure provides a multivariate estimate of the contribution of each region to classifier performance, one should be aware that the maps describe a nonlinear multivariate pattern. Importance scores for each of the regions should be interpreted in the context of the entire multivariate pattern. Against this background, we also present a more readily interpretable, univariate mapping procedure in which we computed classification accuracies for each of the 55 regions separately as described above. A region is shown if the accuracy estimate for this region exceeded chance level (P < .05).

As described above, we corrected for multiple comparisons (for 55 regions) using the false discovery rate.

Results

Regional GPCs
Independent regional GPCs from each of the 55 brain regions revealed significant accuracies for conditioned responses during acquisition (CS+ unpaired > CS−) for the inferior frontal gyrus (pars triangularis), which was greater than chance level (accuracy, 69%; P = .006) but did not survive multiple comparison correction. For the contrast reflecting conditioned responses during extinction (CS+ > CS−), 2 regions were identified: the anterior division of the middle temporal gyrus (accuracy, 65%; P = .03) and the postcentral gyrus (accuracy, 67%; P = .03). Again, none of these displayed accuracy greater than chance after correction for multiple comparisons.

Integration of Regional Predictive Probabilities
Integrating the descriptive probabilities from all regional classifiers for the acquisition and the extinction phases of the fear-conditioning task yielded high predictive performance (accuracy, 82%; sensitivity, 92%; specificity, 72%; P < .001). This finding constituted an improvement in accuracy of 13% (P = .04) compared with the single best of all local classifiers. Data from the acquisition phase alone enabled 73% correct individual-patient classifications (sensitivity, 80%; specificity, 67%; P < .001). For the contrast extinction (CS+ > CS−), integrating the descriptive probabilities from all regional classifiers led to an accuracy of 74% (sensitivity, 64%; specificity, 83%; P < .001).

When we considered the patterns learned by the whole-brain classifier for both contrasts of interest, integrating regional predictions using a GPC algorithm substantially boosted classification accuracy by integrating GPC-predictive probabilities (Figure and Table 2). These regions did not have the highest single GPC accuracies but represent the most informative regions in a multivariate, whole-brain framework for both contrasts combined.

We also conducted a tougher test by taking into account a nonsignificant trend in age differences and a second assessment just before the start of the therapy (Panic and Agoraphobia Scale values28). Because the acquisition of Panic and Agoraphobia Scale values thus technically took place when the treatment had already begun, this analysis might be too conservative. Combining both data sources as described above yielded an accuracy of 79% (sensitivity, 76%; specificity, 82%; P < .001), with an accuracy of 70% (sensitivity, 64%; specificity, 76%; P < .001) for the acquisition phase and 71% (sensitivity, 74%; specificity, 68%; P < .001) for the extinction phase. Although nominally lower than the accuracies reported in the first analyses, the accuracies obtained when taking into account age and Panic and Agoraphobia Scale values did not significantly differ from those obtained in the first analyses (P > .40).

The maps display a region’s contribution to overall classification accuracy combining the acquisition (left) and the extinction (right) phases. Only the highest 10% of the weights are displayed. Colors represent the absolute whole-brain GPC weights; numbers, arbitrary units (AU) to indicate the relative contribution of a region to the prediction of treatment response.

Figure. Multivariate Gaussian Process Classifier (GPC) Weight Maps

Integration of Regional Predictive Probabilities
Integrating the descriptive probabilities from all regional classifiers for the acquisition and the extinction phases of the fear-conditioning task yielded high predictive performance (accuracy, 82%; sensitivity, 92%; specificity, 72%; P < .001). This finding constituted an improvement in accuracy of 13% (P = .04) compared with the single best of all local classifiers. Data from the acquisition phase alone enabled 73% correct individual-patient classifications (sensitivity, 80%; specificity, 67%; P < .001). For the contrast extinction (CS+ > CS−), integrating the descriptive probabilities from all regional classifiers led to an accuracy of 74% (sensitivity, 64%; specificity, 83%; P < .001).

When we considered the patterns learned by the whole-brain classifier for both contrasts of interest, integrating regional predictions using a GPC algorithm substantially boosted classification accuracy by integrating GPC-predictive probabilities (Figure and Table 2). These regions did not have the highest single GPC accuracies but represent the most informative regions in a multivariate, whole-brain framework for both contrasts combined.

We also conducted a tougher test by taking into account a nonsignificant trend in age differences and a second assessment just before the start of the therapy (Panic and Agoraphobia Scale values28). Because the acquisition of Panic and Agoraphobia Scale values thus technically took place when the treatment had already begun, this analysis might be too conservative. Combining both data sources as described above yielded an accuracy of 79% (sensitivity, 76%; specificity, 82%; P < .001), with an accuracy of 70% (sensitivity, 64%; specificity, 76%; P < .001) for the acquisition phase and 71% (sensitivity, 74%; specificity, 68%; P < .001) for the extinction phase. Although nominally lower than the accuracies reported in the first analyses, the accuracies obtained when taking into account age and Panic and Agoraphobia Scale values did not significantly differ from those obtained in the first analyses (P > .40).

The maps display a region’s contribution to overall classification accuracy combining the acquisition (left) and the extinction (right) phases. Only the highest 10% of the weights are displayed. Colors represent the absolute whole-brain GPC weights; numbers, arbitrary units (AU) to indicate the relative contribution of a region to the prediction of treatment response.
Table 2. Top 10% Whole-Brain GPC Weights for the Combined Acquisition and Extinction Contrasts

<table>
<thead>
<tr>
<th>Region of Interest*</th>
<th>Contrastb</th>
<th>Absolute Whole-Brain GPC Weightsc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyrus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precentral</td>
<td>Acquisition: CS+ unpaired &gt; CS−</td>
<td>3.19</td>
</tr>
<tr>
<td>Occipital fusiform</td>
<td>Acquisition: CS+ unpaired &gt; CS−</td>
<td>3.04</td>
</tr>
<tr>
<td>Frontal orbital cortex</td>
<td>Acquisition: CS+ unpaired &gt; CS−</td>
<td>2.79</td>
</tr>
<tr>
<td>Middle temporal gyrus temporo-occipital part</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.78</td>
</tr>
<tr>
<td>Putamen</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.68</td>
</tr>
<tr>
<td>Paracingulate gyrus</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.60</td>
</tr>
<tr>
<td>Supramarginal gyrus</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.47</td>
</tr>
<tr>
<td>Pole</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.23</td>
</tr>
<tr>
<td>Occipital</td>
<td>Extinction: CS+ &gt; CS−</td>
<td>2.15</td>
</tr>
<tr>
<td>Inferior frontal gyrus pars triangularis</td>
<td>Acquisition: CS+ unpaired &gt; CS−</td>
<td>2.15</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>Acquisition: CS+ unpaired &gt; CS−</td>
<td>2.03</td>
</tr>
</tbody>
</table>

Abbreviations: CS, conditioned stimulus; GPC, gaussian process classifier.

Discussion

Although considerable improvements in understanding the neural substrates of anxiety disorders have been achieved, the utility of this knowledge regarding clinical application is limited. As a proof-of-concept study in PD/AG, this report demonstrates that fMRI data acquired before CBT allow for response prediction on an individual-patient level. When we integrated regional predictive probabilities from whole-brain fMRI data, patients could be classified correctly as responders or nonresponders with an accuracy of 82%. Controlling for potential confounders still yielded an accuracy of 79%.

Machine learning has been successfully applied to classify disease state or to predict the onset of disease in neurological or psychiatric patient groups. For psychiatric disorders, classification accuracies based on fMRI data have been reported in the range of 84% to 92% for schizophrenia or from 67% to 86% in major depression. Compared with the relatively large number of studies on disease classification, surprisingly few exist on prognostic markers, almost all of which focus on major depression. With a single exception, investigated all of these studies investigated response to pharmacotherapy. In the only available study on treatment response to CBT, Costafreda et al observed a sensitivity of 71% and a specificity of 86%, leading to an accuracy of 78% for predicting response in a small sample of 16 patients with major depression. In a similar vein, Doehrmann et al demonstrated the potential of fMRI data in predicting treatment response to CBT in social anxiety disorder. Present findings with an accuracy rate of 82% are comparable to the highest obtained in this field. After controlling for potential confounders, the accuracy rate was still in the range of previous reports. Because the Panic and Agoraphobia Scale test was taken within a second assessment just before CBT onset and not after study inclusion, these differences may not reflect true baseline differences, rendering the reanalysis possibly too conservative.

Because our study used a different functional paradigm and aimed at predicting treatment response in patients with PD/AG and not major depression, comparison of regional contributions to overall prediction is of limited utility. Costafreda et al applied an emotional face perception task in patients with major depression, whereas the present study investigated neural responses during fear conditioning as a pathomechanism in PD/AG. Fear conditioning has been reported to activate a widespread network consisting of subcortical structures, such as the thalamus, amygdala, and hippocampus, but also the anterior cingulate cortex, prefrontal and orbitofrontal cortices, and temporal regions. In previous reports on this sample, PD/AG patients exhibited increased inferior frontal gyrus activity during differential conditioning, which possibly indicates aberrant cognitive processing or behavioral inhibition that was sensitive to treatment. The regions contributing to the predictive performance in the present analysis (combined whole-brain acquisition and extinction data) partly overlap with fear network circuits associated with fear conditioning, including the orbitofrontal cortices and the inferior frontal gyrus.

From a health economics viewpoint, identifying potential nonresponders before an intervention is highly desirable. Patients who are not likely to respond could be provided with additional treatment options. Experimental augmentation strategies, such as repetitive transcranial magnetic stimulation, or cognitive enhancers, such as d-cycloserine, are undergoing evaluation for anxiety disorders that include PD. A prognostic marker with high sensitivity would be desirable to stratify patients to available add-on therapies and therefore speed up treatment time and increase response rates. In that sense, a sensitivity of 92% represents a near-maximum detection rate. Patients predicted not to benefit might respond favorably from intensified psychotherapy or neurotherapy/pharmacotherapy, which has yet to be established. Nevertheless, such approaches might aid to govern treatment regimens in the sense of personized medicine approaches. Future research should apply prognostic markers in an independent sample and evaluate the increment in treatment success and cost-effectiveness. We acknowledge that neuroimaging is not yet a routine diagnostic tool in psychiatry. Thus, health economic analyses are needed to balance the costs of additional diagnostic procedures against the potential benefits of improved outcomes.

The present results are based on a medication-free patient sample treated solely with CBT, which limits their generalization to other populations with PD/AG or anxiety disorders and to alternative (eg, pharmacological) treatment. To conform to
Conclusions

We applied a novel machine-learning approach to investigate the potential of fMRI data for CBT response prediction on the individual level in a large sample of PD/AG patients. Findings showed an accuracy of 82% with a high sensitivity of 92%. The present work can be understood as a proof-of-concept study showing the potential utility of our approach for individual-patient response prediction. This multivariate, individual-patient classification approach substantially facilitates current translational efforts aimed at personalizing treatment and supporting clinical decisions when selecting among treatment options, thus helping to improve response rates.
REFERENCES


