Research paper

Neuroimaging, genetic, clinical, and demographic predictors of treatment response in patients with social anxiety disorder

Andreas Frick a,b,⁎, Jonas Engman b, Iman Alaie b,c, Johannes Björkstrand b,d,e, Malin Gingnell b,f, Elna-Marie Larsson g, Elias Eriksson h, Kurt Wahlstedt b, Mats Fredrikson b,i, Tomas Furmark b

a The Beijer Laboratory, Department of Neuroscience, Uppsala University, Uppsala, Sweden
b Department of Psychology, Uppsala University, Uppsala, Sweden
c Department of Neuroscience, Child and Adolescent Psychiatry, Uppsala University, Uppsala, Sweden
d Department of Psychology, University of Southern Denmark, Odense, Denmark
e Department of Psychology, Lund University, Lund, Sweden
f Department of Neuroscience, Uppsala University, Uppsala, Sweden
g Department of Surgical Sciences/Radiology, Uppsala University, Uppsala, Sweden
h Department of Pharmacology, Institute of Neuroscience and Physiology at the Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
i Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

ARTICLE INFO

Keywords:
Social phobia
SSRI
CBT
Personalized medicine
SVM
Pattern recognition

ABSTRACT

Background: Correct prediction of treatment response is a central goal of precision psychiatry. Here, we tested the predictive accuracy of a variety of pre-treatment patient characteristics, including clinical, demographic, molecular genetic, and neuroimaging markers, for treatment response in patients with social anxiety disorder (SAD).

Methods: Forty-seven SAD patients (mean ± SD age 33.9 ± 9.4 years, 24 women) were randomized and commenced 9 weeks' Internet-delivered cognitive behavior therapy (CBT) combined either with the selective serotonin reuptake inhibitor (SSRI) escitalopram (20 mg daily [10 mg first week], SSRI+CBT, n = 24) or placebo (placebo+CBT, n = 23). Treatment responders were defined from the Clinical Global Impression-Improvement scale (CGI-I ≤ 2). Before treatment, patients underwent functional magnetic resonance imaging and the Multi-Source Interference Task taxing cognitive interference. Support vector machines (SVMs) were trained to separate responders from nonresponders based on pre-treatment neural reactivity in the dorsal anterior cingulate cortex (dACC), amygdala, and occipital cortex, as well as molecular genetic, demographic, and clinical data. SVM models were tested using leave-one-subject-out cross-validation.

Results: The best model separated treatment responders (n = 24) from nonresponders based on pre-treatment dACC reactivity (83% accuracy, P = 0.001). Responders had greater pre-treatment dACC reactivity than non-responders especially in the SSRI+CBT group. No other variable was associated with clinical response or added predictive accuracy to the dACC SVM model.

Limitations: Small sample size, especially for genetic analyses. No replication or validation samples were available.

Conclusions: The findings demonstrate that treatment outcome predictions based on neural cingulate activity, at the individual level, outperform genetic, demographic, and clinical variables for medication-assisted Internet-delivered CBT, supporting the use of neuroimaging in precision psychiatry.

1. Introduction

Anxiety disorders, including social anxiety disorder (SAD) (American Psychiatric Association, 2013), are associated with considerable individual suffering and societal costs (Olesen et al., 2012).

Selective serotonin reuptake inhibitors (SSRIs) and cognitive behavior therapy (CBT) are first line treatments for these common and debilitating psychiatric conditions (Baldwin et al., 2014; National Institute for Health and Care Excellence, 2013). Furthermore, in the clinic, it is common to combine SSRIs and CBT (Baldwin et al., 2014;
National Institute for Health and Care Excellence, 2013), and we recently showed that the SSRI escitalopram potentiates the effects of CBT for SAD (Gingnell et al., 2016). Still, only a small majority of patients respond to first line treatment (Baldwin et al., 2014; Taylor et al., 2012) and proper identification of these individuals before treatment initiation would lessen the burden of failed treatment and be a step towards precision psychiatry by enabling more individually tailored interventions. To this end, pre-treatment patient characteristics and neurobiological markers have been used as predictors of interest for treatment outcome (Eskildsen et al., 2010; Frick et al., 2018; Luken et al., 2016; Mululo et al., 2012).

Patient clinical and demographic characteristics associated with better treatment outcome include later age of onset, less severe symptoms, and absence of psychiatric comorbidity (Mululo et al., 2012), but the findings are mixed (Eskildsen et al., 2010). Moreover, anxiety disorders are linked to changes in serotonin neurotransmission, (Durant et al., 2010) including the serotonin transporter (5-HTT) (Frick et al., 2015, 2016b; Maron et al., 2012) and serotonin synthesis (Frick et al., 2016a, 2015; Furmark et al., 2016), and a recent meta-analysis reported the serotonin transporter-linked polymorphic region (5-HTTLPR) to be a potential biomarker for treatment prediction (Luken et al., 2016). The short (s) relative to the long (I) polymorphism of the 5-HTTLPR reduces the transcription of the 5-HTT gene (SLC6A4) (Lesch et al., 1996) and has been associated with enhanced risk for developing excessive anxiety (Gressier et al., 2013) and poorer responses to treatment (Luken et al., 2016). Also, brain serotonin synthesis depends on the tryptophan hydroxylase 2 (TPH2) enzyme and we previously showed that SAD patients carrying a T allele in the TPH2 G-703T single nucleotide polymorphism (SNP) (rs4570625) have heightened serotonin synthesis (Furmark et al., 2016) and are more likely to respond to placebo (Furmark et al., 2008). In addition to serotonergic genes, other genetic polymorphisms have been evaluated in treatment prediction studies, including the brain derived neurotrophic factor (BDNF) val66met SNP (rs66265) and the catechol-o-methyltransferase (COMT) val158met SNP (rs4680) (Luken et al., 2016). However, it should be noted that findings are mixed (Hedman et al., 2012) and that the predictive power of single candidate genes likely is relatively low.

Furthermore, neuroimaging markers of treatment response have attracted a great deal of attention (Luken et al., 2016; Maron and Nutt, 2015; Shin et al., 2013), especially brain reactivity (Doehrmann et al., 2013) and connectivity (Whitfield-Gabrieli et al., 2015). Amygdala reactivity in response to stress is heightened in patients with anxiety disorders as compared to healthy controls (Bührle et al., 2014; Etkin and Wager, 2007; Shin and Liberzon, 2010), and pre-treatment amygdala reactivity has been demonstrated to separate responders from nonresponders to CBT (Bryant et al., 2008; Klumpp et al., 2014; McClure et al., 2007) and SSRI-treatment (McClure et al., 2007) for anxiety disorders. Neural reactivity in the dorsal and ventral occipital cortices has also been used to predict treatment response to CBT for SAD (Doehrmann et al., 2013). Regarding accurate treatment predictions in anxiety disorders, as well as in depression, the anterior cingulate cortex (ACC) is the most consistently reported brain area (Fu et al., 2013; Luken et al., 2016) and specifically the dorsal subdivision of the ACC (dACC) may be a useful predictor of treatment response in SAD (Frick et al., 2018; Klumpp et al., 2016, 2013; Månsson et al., 2015). However, to date, the majority of prediction studies have been limited to monotherapies (Luken et al., 2016) and we only recently reported that pre-treatment dACC response to aversive faces may be a putative marker for decisions on augmenting CBT with SSRI (Frick et al., 2018). Pre-treatment dACC reactivity was higher in responders than in non-responders to combined SSRI + CBT treatment, whereas the opposite was found in the placebo + CBT group.

One drawback of several of the previous treatment prediction studies is the lack of predictions at the individual level (Luken et al., 2016), which is necessary for clinical applications. Supervised machine learning techniques, such as support vector machine (SVM), enable prediction of treatment outcome for individual patients e.g., based on the pattern of brain activity within a region, in contrast to mass-uni-}

2. Methods

The behavioral and neural treatment effects and moderators of treatment outcome have been reported previously for the current sample (Frick et al., 2018; Gingnell et al., 2016), albeit for a different functional magnetic resonance imaging (fMRI) paradigm involving emotional faces. Briefly, SSRI + CBT outperformed placebo + CBT on proportion of responders, 67% and 33% respectively, and symptom improvement was related to attenuated amygdala reactivity to the emotional face-matching task (Gingnell et al., 2016). We here focus on treatment prediction using machine learning (i.e. SVM) and a different fMRI paradigm (the Multi-source interference task; MSIT) (Bush et al., 2003), and also add other predictors.

2.1. Participants

Forty-eight participants (mean ± SD age 33.2 ± 8.8 years, 24 women) meeting the DSM-IV (American Psychiatric Association, 2000) criteria for SAD were randomized to double-blind treatment augmentation of CBT either with escitalopram (SSRI + CBT; n = 24) or pill placebo (placebo + CBT; n = 24). One participant in the placebo + CBT group did not commence treatment and was therefore excluded from the analyses. Thus, 47 participants (mean ± SD age 33.9 ± 9.4 years, 24 women) were included in the study (see Table 1 for participant characteristics and Fig. 1A for a cohort diagram). One additional participant did not partake in post-treatment data collection, leaving 46 participants in the analyses of neural changes with treatment.

Participants were recruited through newspaper advertisements and initially screened for presence of SAD using the Social Phobia Screening Questionnaire (SPSQ) (Furmark et al., 1999) administered online. Participants passing the screening were interviewed using the Mini International Neuropsychiatric Interview (MINI) (Allgulander et al., 2006; Sheehan et al., 1998) and the SAD section from the Structured Clinical Interview for DSM-IV (SCID-I; First et al., 1998). Social anxiety symptom severity was assessed by an experienced psychiatrist (K.W.) using the clinician-administered Liebowitz Social Anxiety Scale (LSAS) (Fresco et al., 2001) and depressive symptoms were measured with the self-report version of the Montgomery Asberg Depressive Rating Scale...
Table 1
Participant characteristics of treatment responders and nonresponders prior to treatment, including statistical tests of group differences.

<table>
<thead>
<tr>
<th>Demographic and clinical information</th>
<th>Responders (n = 24)</th>
<th>Nonresponders (n = 23)</th>
<th>Statistic</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years; mean (SD)</td>
<td>34.3 (10.9)</td>
<td>33.4 (7.8)</td>
<td>t = 0.339</td>
<td>0.736</td>
</tr>
<tr>
<td>Sex, women, n (%)</td>
<td>12 (50%)</td>
<td>12 (52.2%)</td>
<td>χ² = 0.022</td>
<td>0.882</td>
</tr>
<tr>
<td>Symptom severity – LSAS; mean (SD)</td>
<td>70.3 (26.2)</td>
<td>74.6 (22.2)</td>
<td>t = 0.597</td>
<td>0.554</td>
</tr>
<tr>
<td>Depressive symptoms – MADRS-S; mean (SD)</td>
<td>12.5 (7.7)</td>
<td>15.8 (8.0)</td>
<td>t = 1.380</td>
<td>0.175</td>
</tr>
<tr>
<td>Duration of SAD, years; mean (SD)</td>
<td>22.4 (10.3)</td>
<td>24.0 (13.2)</td>
<td>t = 0.458</td>
<td>0.649</td>
</tr>
<tr>
<td>Generalized SAD, n (%)</td>
<td>15 (6.52%)</td>
<td>15 (65.2%)</td>
<td>χ² = 0.236</td>
<td>0.627</td>
</tr>
<tr>
<td>Comorbidity, n (%)</td>
<td>7 (29.2%)</td>
<td>10 (43.5%)</td>
<td>χ² = 1.042</td>
<td>0.307</td>
</tr>
<tr>
<td>Depressive episode, current</td>
<td>2 (8.3%)</td>
<td>1 (4.3%)</td>
<td>–</td>
<td>0.609*</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>4 (16.7%)</td>
<td>6 (25.0%)</td>
<td>–</td>
<td>0.416*</td>
</tr>
<tr>
<td>Obsessive-compulsive disorder</td>
<td>1 (4.2%)</td>
<td>0 (0.0%)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>1 (4.2%)</td>
<td>1 (4.3%)</td>
<td>–</td>
<td>0.609*</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>1 (4.2%)</td>
<td>1 (4.3%)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Earlier psychological treatment, n (%)</td>
<td>3 (12.5%)</td>
<td>1 (4.3%)</td>
<td>–</td>
<td>0.690*</td>
</tr>
<tr>
<td>Earlier pharmacological treatment, n (%)</td>
<td>5 (20.8%)</td>
<td>2 (8.7%)</td>
<td>–</td>
<td>0.416*</td>
</tr>
<tr>
<td>Genetics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-HTTLPR genotype s carriers n (%)</td>
<td>16 (66.7%)</td>
<td>19 (82.6%)</td>
<td>χ² = 0.843</td>
<td>0.358</td>
</tr>
<tr>
<td>TPH2 G-703T genotype T carriers n (%)</td>
<td>13 (54.2%)</td>
<td>13 (59.1%)</td>
<td>χ² = 0.113</td>
<td>0.736</td>
</tr>
<tr>
<td>COMT val158met (rs4680) genotype met carriers n (%)</td>
<td>19 (82.7%)</td>
<td>16 (72.7%)</td>
<td>–</td>
<td>0.491*</td>
</tr>
<tr>
<td>BDNF val66met (rs6265) genotype met carriers n (%)</td>
<td>4 (17.4%)</td>
<td>4 (17.4%)</td>
<td>–</td>
<td>1.000</td>
</tr>
</tbody>
</table>

5-HTTLPR: Serotonin transporter linked polymorphic region; BDNF: Brain-derived neurotrophic factor; COMT: Catechol-O-methyl transferase; LSAS: Liebowitz Social Anxiety Scale; MADRS-S: Montgomery Asberg Depression Rating Scale – Self-Report version; TPH2: Tryptophan hydroxylase 2. a. Fisher’s Exact Test. Cells have an expected count less than 5; b. 22 nonresponders; c. 23 responders.

2.2. Treatments

The treatments have also been described in detail elsewhere (Andersson et al., 2014; Furmark et al., 2009; Gingnell et al., 2016). Briefly, CBT was delivered through a validated therapist-guided Internet-based treatment program (Andersson et al., 2014), based on Clark and Wells’ model of SAD (Clark and Wells, 1995). The program included 9 weekly modules, each containing a reading section and homework assignments: CBT and SAD (module 1), the cognitive model of SAD and cognitive restructuring (modules 2–4), exposure exercises (modules 5–7), and social skills training and relapse prevention (modules 8 and 9). Written feedback on each assignment was provided by a therapist, and the next week’s module was delivered only after the participant completed the homework assignment. Compliance to CBT was assessed by registration of completed modules and homework assignments.

The pharmacological treatment augmentation consisted of escitalopram (20 mg, 10 mg first week). Escitalopram or pill placebo was administered in identical capsules, prepared by APL, Stockholm, Sweden, once daily for nine weeks. Compliance to escitalopram was assessed by analyses of blood metabolites collected at the final fMRI scanning visit (Gingnell et al., 2016).

2.3. Study procedure

Before treatment initiation, participants underwent fMRI with a cognitive test (MSIT) (Bush et al., 2003), and provided a saliva sample for DNA analysis, using an Oragene DNA Genotek saliva sample container. After 9 weeks of treatment, participants again underwent fMRI scanning with the same paradigm, and a blood sample was collected for analyses of escitalopram and metabolite concentrations. An experienced psychiatrist assessed symptom severity before and after 9 weeks of treatment and rated clinical response status using the Clinical Global...
Impression, Improvement scale (CGI-I), at the same post-treatment visit. A CGI-I score of $\leq 2$, corresponding to very much or much improved, was defined as treatment response, and participants scoring $>2$ were defined as non-responders.

2.4. Image acquisition

MRI was performed with a Philips Achieva 3.0T whole body MR scanner (Philips Medical Systems, Best, the Netherlands) equipped with an 8-channel head-coil. Anatomical $T_1$-weighted images were collected using the following settings: echo time (TE):15 ms; repetition time (TR):5700 ms; inversion time:400 ms; field of view:230 $\times$ 230 mm$^2$; voxel size:0.8 $\times$ 1.0 $\times$ 2.0 mm$^3$; 60 contiguous slices. Functional MRI image acquisition utilized a blood oxygenation level-dependent (BOLD) echo planar imaging (EPI) sequence (TE:35 ms; TR:3000 ms; slice thickness 4 mm; 64 axial slices). Acquisition parameters from the realignment step (to remove movement-related confounds) were defined and the statistical threshold set to P<0.05 family-wise error (FWE) correction.

2.5. FMRI task

Participants underwent the cognitive interference task MSIT (Bush et al., 2003) during the fMRI image acquisition (Fig. 1B). The task had 4 blocks each of the two conditions (control and interference), and fixation crosses were presented for the first and last 30 s of image acquisition, for a total length of 6 min and 36 s. Each block was 42 s and consisted of 24 trials (1750 ms per trial) in a pseudo-randomized order with sets of 3 digits (0, 1, 2, or 3) displayed in the center of the display. One digit (target) was different from the other two (distractors). In the control condition, the distractors were always ‘0’ and the target digit corresponded to its position (i.e. ‘1’ was always in the leftmost position 1; ‘2’ in the middle position 2, middle; and ‘3’ in the rightmost position 3). Thus, in the control condition, the target digit and position were congruent. In the interference condition, distractors were either ‘1’, ‘2’, or ‘3’ and the target digit and position were incongruent. The participants indicated if the target digit was ‘1’, ‘2’, or ‘3’ with the response buttons (‘1’: left thumb, ‘2’: left index finger, ‘3’: right index finger). Interference blocks were contrasted against meters from the realignment step (to remove movement-related confounds).

2.6. Image analyses

Standardized preprocessing steps for fMRI images were conducted in Statistical Parametric Mapping 8 (SPM8; www.fil.ion.ucl.ac.uk/spm/software/spm8) inside MATLAB (MathWorks, Natick, Massachusetts, USA). Each participant’s BOLD EPI images were realigned to the mean EPI image, slice timing corrected to the middle slice of each volume, and co-registered to the anatomical $T_1$-weighted image. Normalization parameters from Unified segmentation of the anatomical image were used to normalize the EPI images to Montreal Neurological Institute (MNI) standard space. Subsequent spatial smoothing was applied with an 8 mm Gaussian kernel. None of the participants had movement exceeding our threshold for exclusion, i.e. one voxel (3 mm).

First-level analyses consisted of high-pass filtering (128 s) and general linear modeling of the stimulus function (boxcar, onsets and durations of control and interference blocks) and six movement parameters from the realignment step (to remove movement-related components of the data) convolved with the canonical haemodynamic response function in SPM8. Interference blocks were contrasted against control blocks and used in subsequent analyses.

2.7. Genotyping

Genotyping of the 5-HTTLPR, G-703T polymorphism (rs4570625) in the promoter of the TPH2 gene, the val158met polymorphism (rs4680) in the gene encoding COMT, and the val66met polymorphism (rs6265) in the BDNF gene, were performed from saliva samples (see Supplementary information). TPH2 gene analyses failed for one participant, and COMT analyses failed for two other participants, leaving 44 participants with genotype information for the 4 included polymorphisms.

2.8. Statistical analyses

2.8.1. SVM prediction of treatment response

SVM analyses were carried out in R 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria) using the package e1071 (cran.r-project.org/web/packages/e1071). Treatment responders and nonresponders were entered as two classes, weighted by number of participants, in linear kernel SVMs with scaled data and the SVM soft margin parameter C fixed to its default value (one). Leave-one-subject-out cross-validations were performed to assess generalization of prediction accuracies. The statistical significance of the classifications was tested using permutation testing with 1000 permutations with random assignment of group class to input image. The resulting null-hypothesis distribution was used to calculate the p-value of the accuracies, i.e. the proportion of permutations that yielded a greater or equal accuracy than the accuracy found for the classification models. P<0.05 was used indicative of statistical significance.

We constructed 6 a priori SVM models with treatment response as outcome. SVM model 1 (dACC model) included each individual’s extracted contrast values (interference – control) from the cognitive interference task from all voxels (i.e. no threshold was used) in the dACC ROI, defined as the portion of the cingulate cortex between $Y = 0$ and $Y = 30$ mm and superior to $Z = 0$ mm (Bush et al., 2003). SVM model 2 (amygdala model) included extracted contrast values from the bilateral amygdala, as defined in the Automatic Anatomical Labeling Atlas (AAL) (Maldjian et al., 2003). SVM models 3 and 4 (dOCC model and vOCC model) included extracted contrast values from 10 mm radius spheres around the peak voxel from the dorsal and ventral occipital cortex clusters, respectively, found to be predictive of response to CBT treatment by Doehrmann et al. (2013). SVM model 5 (demographic and clinical model) included demographic and clinical characteristics, i.e. age, sex, symptom severity (LSAS), subtype of SAD (generalized or specific), comorbidity status, previous psychological treatment, previous pharmacological treatment, and depression score (MADRS-S). SVM model 6 (genetic model) included the following genotype information: s (5-HTTLPR), T (TPH2 G-703T), met (COMT val158met) and met (BDNF val66met) carrier status.

In addition to the a priori SVM models, we added demographic, clinical, and genetic information respectively, to any significant brain-based model.

2.8.2. Pre-treatment differences between responder groups

To complement the SVM analyses and test the association between treatment response and pre-treatment neural reactivity at the group level, analyses of the direction of the differences between treatment responders and nonresponders in pre-treatment neural reactivity to cognitive interference were carried out with two-sample t-tests in SPM8. ROI analyses, corresponding to the SVM models 1–4, were performed and the statistical threshold set to P<0.05 family-wise error corrected for multiple comparisons within the ROI (P_{FWE}). Additionally, exploratory whole-brain analyses were performed with the statistical threshold set to P_{FWE}<0.05.

Differences in demographic, clinical, and genetic characteristics were tested using two-sample t-tests, chi-square tests, and Fisher exact tests in R 3.1.1. P<0.05 indicated statistical significance.

2.8.3. Changes in neural reactivity with treatment

To test if neural reactivity changed with treatment within any ROI with significant outcome predictions, we created pre-post difference images with the SPM module ImCalc and conducted two-sample t-tests between responder groups in SPM8.
2.9. Ethics statement

The study was approved by the Regional Ethical Review Board, Uppsala and the Medical Products Agency in Sweden. All participants were fully informed about the study aims and procedures and gave written informed consent prior to inclusion.

3. Results

3.1. SVM prediction of treatment response

Pre-treatment dACC reactivity to cognitive interference (SVM model 1) accurately separated the 24 responders from the 23 nonresponders (balanced accuracy 83%, \( P = 0.001 \); sensitivity 87.5%, specificity 78.3%). When we applied the model created on the whole sample to the treatment groups separately, number of accurately classified patients did not differ between the SSRI+CBT (83.3% accuracy) and placebo+CBT (82.6% accuracy) groups (\( \chi^2 = 0, P = 1 \)). A follow-up support vector regression analysis with pre-to-post change on LSAS (mean ± SD change \(-27.8 \pm 20.9\)) as outcome showed that pre-treatment dACC reactivity also was predictive of continuous symptom improvement (\( r = 0.44, P = 0.013 \)). Amygdala, dOCC or vOCC reactivity, demographic and clinical variables, or genetic information models did not produce accurate classifications of responder status, and the addition of demographic and clinical data to the dACC SVM model tended to worsen classifier performance (Table 2).

The results remained when considering only patients who completed the whole CBT treatment (i.e., all 9 weekly modules; Supplementary Table 1). Only the dACC SVM model provided significant predictive accuracy (81.3%, \( P = 0.008 \)).

3.2. Pre-treatment differences between responder groups

Voxel-wise ROI analyses revealed higher pre-treatment dACC reactivity in responders relative to nonresponders, albeit only at trend-level significance (MNI \( x, y, z = 12, 23, 38 \); \( Z = 3.12, P_{FWE} = 0.087, P_{uncorrected} = 0.001 \)) (Fig. 2). The difference in pre-treatment reactivity between treatment responders and nonresponders survived stringent statistical thresholding when restricting the analyses to those treated with combined SSRI+CBT (MNI \( x, y, z = 12, 20, 34 \); \( Z = 3.54, P_{FWE} = 0.032, 135 \text{ mm}^3 \)), but not when considering the placebo+CBT group only (\( P_{FWE} > 0.7 \)). At a liberal statistical threshold (\( P_{uncorrected} < 0.05 \)), the directions of findings in the placebo+CBT group were mixed, suggesting no clear relation between pre-treatment dACC activity and placebo+CBT treatment outcome. In exploratory whole-brain analyses, no pre-treatment differences were noted between responders and nonresponders.

Treatment responders did not differ from nonresponders on any of the demographic, clinical, or molecular genetic variables in the whole sample or in any of the treatment groups separately (Table 1, Supplementary Tables 2 and 3).

Table 2

<table>
<thead>
<tr>
<th>SVM Model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. dACC</td>
<td>83.0% (39/47)</td>
<td>87.5% (21/24)</td>
<td>78.3% (18/23)</td>
</tr>
<tr>
<td>2. Amygdala</td>
<td>38.3% (18/47)</td>
<td>37.5% (9/24)</td>
<td>59.1% (9/23)</td>
</tr>
<tr>
<td>3. dOCC</td>
<td>53.2% (25/47)</td>
<td>58.3% (14/24)</td>
<td>47.8% (11/23)</td>
</tr>
<tr>
<td>4. vOCC</td>
<td>36.2% (17/47)</td>
<td>41.7% (10/24)</td>
<td>30.4% (7/23)</td>
</tr>
<tr>
<td>5. Demographic and clinical</td>
<td>48.9% (23/47)</td>
<td>66.7% (16/24)</td>
<td>30.4% (7/23)</td>
</tr>
<tr>
<td>6. Geneticaa</td>
<td>15.9% (7/44)</td>
<td>30.4% (7/23)</td>
<td>0% (0/21)</td>
</tr>
<tr>
<td>7. dACC, demographic and clinical</td>
<td>68.1% (32/47)</td>
<td>75.0% (18/24)</td>
<td>60.9% (14/23)</td>
</tr>
<tr>
<td>8. dACC and genetica</td>
<td>61.4% (27/44)</td>
<td>65.2% (15/23)</td>
<td>57.1% (12/21)</td>
</tr>
</tbody>
</table>

dACC: dorsal anterior cingulate cortex; dOCC: dorsal occipital cortex; vOCC: ventral occipital cortex.

a. \( n = 44 \).
3.3. Changes in neural reactivity with treatment

No differences between responders and nonresponders in neural reactivity changes were detected within the dACC in the ROI analysis ($P_{FWE} > 0.188$) or in whole-brain exploratory analyses ($P_{FWE} > 0.745$).

4. Discussion

The results of this machine learning fMRI study converge with previous findings of more accurate predictions from neural than demographic and clinical variables (Ball et al., 2014; Doehrmann et al., 2013; Frick et al., 2018; Månsson et al., 2015; Whitfield-Gabrieli et al., 2015), and more specifically that the dACC (Klumpp et al., 2016, 2013; Månsson et al., 2015), but not patient clinical/demographic characteristics (Eskildsen et al., 2010) are predictive of SAD patients’ response to CBT, both when delivered face-to-face and via the Internet as in this study, and when combined with an SSRI. The SVM model based on dACC reactivity to cognitive interference achieved 83% accurate predictions, which is similar to a recent study using SVM and pre-treatment dACC reactivity to self-referential criticism to predict CBT treatment outcome at one year follow-up (Månsson et al., 2015). Thus, dACC reactivity has the potential to improve predictions of treatment outcome and clinical decision-making.

The accurate predictions of treatment outcome from dACC reactivity to cognitive interference in the present study add to previous findings of a relationship between dACC reactivity to emotional stimuli and clinically relevant symptom improvement following CBT for SAD (Frick et al., 2018; Klumpp et al., 2016, 2013; Månsson et al., 2015). However, the direction of dACC reactivity associated with favorable treatment outcome differs between studies. In the present study, responders relative to nonresponders had increased dACC reactivity to cognitive interference, in line with findings on CBT response from Klumpp et al. who used fearful and angry faces as signals of threat or as distractors during high perceptual load (2016, 2013), whereas Månsson et al. (2015) and Frick et al. (2018) recently reported that CBT responders had lower dACC reactivity than nonresponders to self-referential criticism and aversive faces respectively. It should be noted that the response-related difference in pre-treatment dACC reactivity was restricted to combined SSRI+CBT treatment and not detected in the present placebo+CBT group, in contrast to our previous study using aversive faces in the same sample (Frick et al., 2018). This discrepancy might be related to the use of a cognitive task specifically targeting the dACC (Bush and Shin, 2006) vs. emotional tasks. On the other hand, the dACC SVM model constructed from the whole sample had similar accuracies for both treatment groups (83%), suggesting that pre-treatment dACC reactivity to a cognitive task actually may hold predictive information also for the patients treated with placebo+CBT.

Despite the accurate predictions of response from pre-treatment dACC reactivity, we detected no robust relationship between reduced dACC reactivity following treatment and clinical response. This dissociation is consistent with the notion that pre-treatment predictive information does not necessarily inform on treatment mechanisms (Lueken and Hahn, 2016). Although predictions based on dACC reactivity have the potential to be clinically valuable despite being agnostic to treatment mechanisms, the specific functions subserved by the dACC accounting for the accurate predictions remain to be explained. The cognitive interference paradigm used in the current study was designed specifically to evoke neural activity, induced by cognitive, but not emotional processes, in the dACC and included components such as target/error detection, response selection, and conflict monitoring (Bush et al., 2002; Bush and Shin, 2006). Such cognitive functions may indeed be deficient in individuals with anxiety disorders and be targeted by CBT, although it should be noted they were not directly targeted in the version of CBT used in the present study. In line with our findings of increased pre-treatment dACC reactivity to cognitive conflict in SSRI+CBT responders, Klumpp et al. reported that symptom improvement following CBT for SAD was associated with elevated pre-treatment dACC reactivity to emotional conflict (Klumpp et al., 2016). These findings suggest that predictive dACC reactivity may be related to cognitive and emotional conflict detection and resolution, functions possibly involved in emotion regulation through cognitive reappraisal of emotions also observed by the dACC (Messina et al., 2015). Notably, in patients with depression, increased ACC reactivity (Fu et al., 2013) and pre-treatment performance on cognitive and emotional tests have been related to treatment outcome (Etkin et al., 2015), the latter potentially being mediated by the dACC, but this remains to be tested. Moreover, the dACC is a central node in the salience network, which is involved in detecting relevant internal and external stimuli (Menon, 2015), and anxiety disorders are associated with aberrant attention to internal and external stimuli as well as alterations in the salience network (Menon, 2015). Further studies discerning the dACC functions underlying treatment prediction accuracy are needed, but for now we can conclude that dACC reactivity, be it during cognitive or emotional tasks, is indicative of treatment outcome (Klumpp et al., 2016, 2013; Lueken et al., 2016; Månsson et al., 2015).

In line with previous findings (Hedman et al., 2012), a greater proportion of treatment responders compared to nonresponders completed all of the CBT modules. However, when considering only the patients that completed all CBT modules, the dACC SVM model still accurately predicted treatment response. Notably, all patients complied with SSRI treatment. Thus, differences in treatment compliance do not account for the accurate outcome predictions seen in this study.

Demographic, clinical, and genetic variables did not differ between treatment responders and nonresponders, nor could they accurately predict treatment outcome. Furthermore, adding these variables to the dACC SVM models tended to worsen rather than increase prediction accuracy. It has been reported that less severe symptoms at pre-treatment is associated with a better treatment outcome (Mululo et al., 2012), but this was not confirmed in the present study. In fact, previous findings from prediction studies based on patient clinical and demographic characteristics are mixed (Eskildsen et al., 2010; Mululo et al., 2012). Moreover, in their systematic review, Lueken et al. (2016) reported that the most consistent therapy-genetics finding is the association between the short variant of the 5-HTTLPR and nonresponse to treatment, albeit this was only seen in about one-half of the studies assessing this relation. In the current sample, we could not detect any relationship between genetic markers and treatment outcome, neither in the group level analyses nor in the individual patient prediction analyses.

The current study has some limitations that deserve mentioning. First, the sample size is small, especially for genetic and demographic/clinical prediction studies (Lueken et al., 2016), which makes it prone to type I and II errors and thus may have precluded findings of small effects. This and the relatively large number of tests performed hence calls for replication in independent samples before drawing strong conclusions. Second, we only had access to the diallelic division of 5-HTTLPR genotypes and not to the triallelic 5-HTTLPR/rs25531 LA/LG/S division. Nonetheless, a previous treatment prediction study showed similar accuracies for the diallelic and triallelic divisions of the SHTTLPR (Stein et al., 2006). Furthermore, although leave-one-subject-out cross-validation mitigates the inherent bias in constructing and testing the statistical model on the same sample, while making the most use of a limited sample size, generalizability to independent samples and different scanner sites needs to be addressed in future studies to enable clinical application. We also lacked a group receiving SSRI monotherapy, which would be useful in further understanding the predictive role of dACC reactivity across treatments.

In conclusion, this machine learning study showed that pre-treatment neural reactivity in the dorsal ACC is predictive of outcome in patients with SAD treated with Internet-delivered CBT alone or in combination with an SSRI. In contrast, amygdala, dOCC, and vOCC reactivity as well as demographic, clinical, and genetic information did
not separate responders from nonresponders, and did not add predictive power to the DACC models. Thus, brain-based predictions of treatment outcome outperform demographic, clinical, and genetic variables for individual SAD patients. Successful prediction of treatment response and non-response on an individual level may inform clinicians on treatment selection for individual patients, an important aspect of precision psychiatry.

Role of the funding source

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript, or decision to submit the manuscript for publication.

Contributors

AF, JE, MF, and TF designed the study. AF, JE, IA, JB, MG, KW, MF, and TF collected the data. JE, JB, KW, and TF treated the patients. AF, EML, EE, MF, and TF performed the analyses, and AF and TF wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Funding

The study was supported by the Swedish Research Council, the Swedish Brain Foundation, Riksbanksens Jubileumsfond – the Swedish Foundation for Humanities and Social Sciences, the Swedish Research Council for Health, Working Life and Welfare, and the Kjell and Märta Beijer Foundation.

Declaration of Competing Interest

None.

Acknowledgements

We would like to express our gratitude to all study participants.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2019.10.027.

References


