

## ORIGINAL ARTICLE

# Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning

KNT Månsson<sup>1</sup>, A Frick<sup>2</sup>, C-J Boraxbekk<sup>3,4</sup>, AF Marquand<sup>5,6</sup>, SCR Williams<sup>6</sup>, P Carlbring<sup>7</sup>, G Andersson<sup>1,8</sup> and T Furmark<sup>2</sup>

Cognitive behavior therapy (CBT) is an effective treatment for social anxiety disorder (SAD), but many patients do not respond sufficiently and a substantial proportion relapse after treatment has ended. Predicting an individual's long-term clinical response therefore remains an important challenge. This study aimed at assessing neural predictors of long-term treatment outcome in participants with SAD 1 year after completion of Internet-delivered CBT (iCBT). Twenty-six participants diagnosed with SAD underwent iCBT including attention bias modification for a total of 13 weeks. Support vector machines (SVMs), a supervised pattern recognition method allowing predictions at the individual level, were trained to separate long-term treatment responders from nonresponders based on blood oxygen level-dependent (BOLD) responses to self-referential criticism. The Clinical Global Impression-Improvement scale was the main instrument to determine treatment response at the 1-year follow-up. Results showed that the proportion of long-term responders was 52% (12/23). From multivariate BOLD responses in the dorsal anterior cingulate cortex (dACC) together with the amygdala, we were able to predict long-term response rate of iCBT with an accuracy of 92% (confidence interval 95% 73.2–97.6). This activation pattern was, however, not predictive of improvement in the continuous Liebowitz Social Anxiety Scale—Self-report version. Follow-up psychophysiological interaction analyses revealed that lower dACC–amygdala coupling was associated with better long-term treatment response. Thus, BOLD response patterns in the fear-expressing dACC–amygdala regions were highly predictive of long-term treatment outcome of iCBT, and the initial coupling between these regions differentiated long-term responders from nonresponders. The SVM-neuroimaging approach could be of particular clinical value as it allows for accurate prediction of treatment outcome at the level of the individual.

*Translational Psychiatry* (2015) 5, e530; doi:10.1038/tp.2015.22; published online 17 March 2015

## INTRODUCTION

Social anxiety disorder (SAD) is a common<sup>1</sup> and disabling disorder that often precedes other serious mental health problems such as depression.<sup>2</sup> SAD is associated with aberrant information-processing and cognitive biases toward negative information regarding the self, for example, self-focused attention.<sup>3</sup> Cognitive behavior therapy (CBT), including modification of cognitive biases, is an effective treatment for SAD,<sup>4</sup> and can be delivered in accessible formats, for example, via the Internet.<sup>5</sup> Internet-delivered CBT (iCBT) has been evaluated in several randomized controlled studies at different sites,<sup>5</sup> and sustained effects have been observed up to 5 years later.<sup>6</sup> Also, CBT for SAD has been shown to be equally effective when delivered via the Internet in comparison with face-to-face group treatment.<sup>7</sup> However, significant proportions of the treated patients relapse over time or do not respond sufficiently.<sup>8</sup> Accordingly, improving long-term treatment outcome of CBT remains an important challenge, and factors that reliably predict lasting therapeutic success need to be identified.

Recent functional magnetic resonance imaging (fMRI) studies suggest that neural biomarkers add substantial value to

predictions of CBT outcome. In SAD participants, Doehrmann *et al.*<sup>9</sup> demonstrated that initial activations of the visual cortex, in response to emotional face stimuli, predicted symptom improvement with CBT, and that brain measures vastly improved prediction success in comparison with other clinical variables. Klumpp *et al.*,<sup>10</sup> using a similar task, found that pretreatment reactivity in the inferior frontal and the superior and middle temporal gyri was associated with reduced social anxiety following CBT. However, both studies evaluated the short-term treatment response only. In a previous positron emission tomography study from our lab, the initial attenuation (pre–post) of anxiety-related amygdala activity in treated SAD participants was associated with clinical improvement 1 year later, but as the sample was small the study could not discriminate properly between the effects of CBT and selective serotonin reuptake inhibitors (SSRIs).<sup>11</sup>

Mass-univariate voxel-wise methods, seeking pretreatment brain voxels that correlate with symptom improvement or differ between responders and nonresponders, have so far been the most common approach in psychiatric prediction studies. However, such studies may not be easily translated to clinical settings

<sup>1</sup>Division of Psychology, Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden; <sup>2</sup>Department of Psychology, Uppsala University, Uppsala, Sweden; <sup>3</sup>Centre for Population Studies, Ageing and Living Conditions, Umeå University, Umeå, Sweden; <sup>4</sup>Umeå Center for Functional Brain Imaging (UFBI), Umeå University, Umeå, Sweden; <sup>5</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; <sup>6</sup>Department of Neuroimaging, Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College London, London, UK; <sup>7</sup>Department of Psychology, Stockholm University, Stockholm, Sweden and <sup>8</sup>Psychiatry Section, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. Correspondence: KNT Månsson, MSc, Division of Psychology, Department of Behavioural Sciences and Learning, Linköping University, Linköping SE-581 83, Sweden.

E-mail: kristoffer.nt.mansson@liu.se

Received 19 August 2014; revised 13 January 2015; accepted 20 January 2015

**Table 1.** Demographic and clinical characteristics of study participants

	All participants (n = 26)	Responders at 1 year (n = 12)	Nonresponders at 1 year (n = 11)	Responders vs nonresponders at 1 year (n = 12/11)
Age (years), mean (s.d.)	32.3 (9.6)	35.5 (8.5)	31.5 (10.3)	$t(21) = 1.01, P = 0.837$
Range (years)	19–57	21–47	20–57	
Gender, female (%)	22 (85)	9 (75)	10 (91)	$\chi^2(1) = 1.01, P = 0.315$
Married or <i>de facto</i> , n (%)	15 (58)	7 (58)	7 (64)	$\chi^2(1) = 0.07, P = 0.795$
Educational level, n (%)				Fisher's exact $P = 0.386$
Completed university	9 (35)	5 (42)	3 (27)	
Current university	10 (38)	5 (42)	3 (27)	
Lower grade <sup>a</sup>	7 (27)	2 (17)	5 (45)	
Psychotropic medication, n (%)	8 (31)	3 (25)	4 (36)	$\chi^2(1) = 0.35, P = 0.554$
Age of SAD onset (years), mean (s.d.)	15.9 (6.0)	16.3 (4.2)	16.6 (7.7)	$t(21) = 0.12, P = 0.907$
Pretreatment LSAS-SR, mean (s.d.)	76.3 (18.7)	74.1 (15.1)	77.6 (23.9)	$t(21) = 0.43, P = 0.672$
Pretreatment MADRS-S, mean (s.d.)	15.8 (6.6)	15.4 (8.1)	15.4 (5.8)	$t(21) = -0.02, P = 0.986$

Abbreviations: LSAS-SR, Liebowitz Social Anxiety Scale—Self-report; MADRS-S, Montgomery Åsberg Depression Rating Scale—Self-report; SAD, social anxiety disorder. <sup>a</sup>Including high school, vocational school and compulsory school.

when the concern is whether or not a certain patient will respond to a specific treatment. In contrast, supervised pattern recognition methods constitute a novel approach in clinical neuroimaging, utilizing patterns of information across many voxels, for example, to separate responders from nonresponders.<sup>12</sup> In pattern recognition, a statistical model is estimated on a subset of the data ('training') and then applied to predict the diagnostic or prognostic label of novel unseen data ('testing'). This is usually achieved within a cross-validation procedure, whereby the model is repeatedly retrained, withholding a different partition of data each time. This approach is standard within the statistical literature and is well known to provide approximately unbiased estimates of generalizability to new samples, and helps to minimize the possibility of 'over-fitting' the data. Of crucial importance, and in contrast to voxel-wise group statistics, pattern recognition analyses can make predictions at the level of the individual based on the pattern within the data, for example, treatment responder status or diagnostic category for a new unseen participant.<sup>12,13</sup>

A supervised pattern recognition method called support vector machine (SVM)<sup>13</sup> was recently successful in separating SAD from panic disorder participants, and from healthy controls.<sup>14</sup> In addition, we recently demonstrated the utility of SVM and accurately discriminated SAD participants from healthy controls based on the multivariate pattern of blood oxygen level-dependent (BOLD) response in the fear network,<sup>15</sup> that is, the amygdala, hippocampus, anterior cingulate cortex (ACC) and insula, frequently demonstrated to be dysfunctional in SAD.<sup>16</sup> Pretreatment reactivity in fear network regions has also been associated with subsequent treatment outcome in SAD and other anxiety disorders.<sup>17,18</sup> The ACC in particular has been implicated in treatment prediction studies of depression,<sup>19</sup> and anxiety.<sup>20</sup> Moreover, cognitive control of emotion, involving prefrontal- limbic interactions,<sup>21</sup> has been suggested to be impaired in participants with SAD and restored with CBT<sup>22</sup> but it is not known whether such interactions are predictive of treatment response in SAD. Although the SVM approach has been successful in predicting treatment response in depression,<sup>23</sup> there are, to our knowledge, no studies using SVM to predict CBT treatment outcome in anxiety disorders such as SAD. Also, neuroimaging prediction studies of long-term treatment effects are largely lacking in psychiatric research.<sup>24</sup>

The objective of the present study was to use SVM classification to evaluate neural predictors of long-term iCBT response 1 year after treatment of participants with SAD. In addition to iCBT, the

participants underwent Internet-delivered attention bias modification (ABM) in a cross-over design.<sup>25</sup> Adding ABM has not been shown to further improve outcome,<sup>26</sup> and we refer to the present combined intervention as iCBT. Using a disorder-relevant fMRI paradigm, we entered BOLD responses to sentences with negative content about oneself, that is, self-referential criticism<sup>27</sup> into linear SVMs to classify responder status at long-term follow-up. Our SVM analyses focused primarily on the fear network and brain regions associated with cognitive control of negative emotion, that is, the ACC, amygdala, hippocampus, insula, ventromedial prefrontal cortex and dorsolateral prefrontal cortex.

## MATERIALS AND METHODS

### Participants

This study included 26 right-handed participants with a primary diagnosis of SAD (85% having the generalized subtype, recruited via media advertisements (see Table 1 and Supplementary Figure S1 in Supplementary Material). The sample, including inclusion/exclusion criteria, has been described in detail elsewhere,<sup>25</sup> and the screening procedure was similar to our previous treatment studies.<sup>5,28</sup> Briefly, participants reported interest on a webpage and answered self-report questionnaires via the Internet regarding social anxiety, depression and magnetic resonance safety. Participants had no neurological or major somatic disorder, no suicidal ideation, no other ongoing psychological treatment and they were not included if psychotropic medication (for example, SSRIs) was recently initiated or changed, that is, a stable dose for 3 months was required. Applicants fulfilling the initial screening criteria were interviewed via telephone using the structured clinical interview for the 4th version of Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) axis 1 (SCID-I).<sup>29</sup> At baseline, 8 (31%, 8/26) participants were currently on prescription medication, that is, SSRIs. Five participants had a history of SSRI treatment and 13 were medication-naïve. One participant (deemed as nonresponder) increased the dose from 50–100 mg sertraline from posttreatment to 1-year follow-up. Three participants withdrew and did not take part in the 1-year follow-up.

All participants gave written informed consent prior to participation. The study was conducted in accordance with the Declaration of Helsinki and approval was obtained from the regional ethic committee. The study was registered at ClinicalTrials.gov (ID: NCT01312571).

### Assessment of clinical response

Treatment response at 1-year follow-up was assessed with the Clinical Global Impression-Improvement (CGI-I) interview scale.<sup>30</sup> In accordance with prior studies,<sup>30</sup> scores of 1 or 2 (very much or much improved) defined treatment responders, whereas participants scoring  $\geq 3$  (ranging from minimally improved to very much worse) were classified as

nonresponders. Two psychologists, blind to the experimental conditions, conducted the interviews. In addition, we predicted long-term clinical improvement (pre follow-up) using the continuous Liebowitz Social Anxiety Scale—Self-report version, LSAS-SR.<sup>31</sup>

## Treatment

This study included two forms of intervention for SAD. Both treatments were provided to all participants in a cross-over design where half of the participants started with iCBT and the other half started with ABM.<sup>25</sup> The guided iCBT protocol contained a 9-week intervention supported by a clinician and has been found effective in several randomized controlled trials.<sup>5,28,32,33</sup> Previously, we showed that the short-term outcome in this sample was in favor of iCBT, that is, 66% of the iCBT participants, in comparison with 25% of the ABM group, were deemed as CGI-I responders.<sup>25</sup> At the 1-year follow-up, we found no effect of the order ( $\chi^2=0.4$ ,  $P=0.51$ ) in which the two interventions were presented.

Briefly, the treatment program is based upon the cognitive model of SAD described by Clark and Wells.<sup>3</sup> The self-help text included information, exercises, homework assignments and ended with essay questions that were sent to the therapist. For more details, see Andersson *et al.*<sup>5</sup> and Supplementary Material. The second intervention was ABM delivered over 4 weeks via the Internet, with exercises implemented twice a week, totaling eight sessions. ABM aims at changing the distorted attentional process characterizing anxiety disorders, that is, hypervigilance, or problem of disengagement from fearful stimuli.<sup>34</sup> The Internet-delivered ABM used in this study was a web-based flash program displayed in full-screen mode at the participant's computer. E-mail and mobile phone SMS reminders prompted for each training session, scheduled on Tuesdays and Thursdays over the 4-week treatment course. The participant and therapist had no other communication during the treatment period.

## Experimental task

We evaluated BOLD responses to self-referential criticism using a disorder-relevant paradigm developed by Blair *et al.*,<sup>27</sup> which was translated into Swedish. The sentences were split into three categories based on the referential target: self, other female or other male. The task included 216 sentences divided into three categories based on valence: negative (for example, 'Nobody likes you'), neutral (for example, 'You read') or positive (for example, 'Everyone loves you') valences. Across valences, sentences were matched on number of letters and words, and appeared in randomized order. Participants were instructed to read the sentences and press a button with the right hand when they had read each sentence. Each sentence was presented for a maximum of 2500 ms. In addition, 96 fixation crosses ('+') were randomly interspersed between the sentences and displayed for 2500 ms. Each sentence and fixation cross was separated by a cross or circle presented for 500 ms. Stimuli were demonstrated using the E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, USA), projected on a screen and viewed through a tilted mirror attached to the head coil. The total duration of the task was 17 min and 20 s.

## Imaging

**Structural and functional image acquisition.** Structural high-resolution T1-weighted images (180 slices, 1 mm thickness, field of view: 250 mm, voxel size:  $0.5 \times 0.5 \times 1 \text{ mm}^3$ ) were collected prior to the functional images. BOLD contrast images (T2\* weighted) were acquired using a 3T Discovery MR750 (General Electric, Madison, WI, USA) scanner equipped with a 32-channel head coil. The following scanning parameters were used: echo time: 30 ms, repetition time: 2000 ms, flip-angle:  $80^\circ$ , field of view:  $25 \times 25 \text{ cm}^2$ , matrix size:  $96 \times 96$ . Thirty-seven slices with a thickness of 3.4 mm ( $2.6 \times 2.6 \times 3.4$  voxel size) were acquired every 2000 ms. Ten dummy scans were run before the image acquisition started to avoid signals resulting from progressive saturation. Preprocessing was done using Statistical Parametric Mapping Software 8 (SPM8; Wellcome Department of Cognitive Neurology, London, UK) implemented in MATLAB (Mathworks, Natick, MA, USA). Functional scans were realigned to the mean image of each run to correct for head movements during image acquisition, after which slice timing correction was performed. The functional scans ( $3 \times 3 \times 3 \text{ mm}^3$  voxel dimensions) were then co-registered to the structural scans. Structural scans were segmented into gray and white matter, and parameters for warping the scans to the Montreal Neurological Institute (MNI152) template were calculated using the unified segmentation procedure in SPM8. Functional scans were subsequently warped to MNI standard space by applying these parameters, and smoothed with an 8-mm isotropic Gaussian kernel.

## Data analysis

Functional brain imaging data were fitted to the general linear model using SPM8. First-level, within-subject, analyses included nine regressors of interest in the model, valence ( $3$ )  $\times$  target ( $3$ ), containing onset times of respective sentences, as well as the *isx* movement parameters from the realignment preprocessing step as nuisance regressors. The model was convolved with the canonical hemodynamic response function defined in SPM8 and filtered using a 128-s high-pass filter. The other-referred contrast consisted of both female and male targets, and the contrast of interest used in subsequent analyses was self-referential criticism (self-negative minus other negative). On the basis of previous CBT prediction studies,<sup>11,17,18</sup> we selected fear network regions of interest (ROIs) including the ACC, amygdala, hippocampus and insula. All regions were defined by the Automatic Anatomical Labeling ROI library within the Wake Forest University PickAtlas software.<sup>35</sup> In follow-up analyses, the ACC was divided into the dorsal (dACC) and ventral (vACC) subdivisions by a split at the MNI-coordinate  $z=8$ , which corresponds to clusters 2 and 3 in Beckmann *et al.*<sup>36</sup>

We also investigated the predictive value of regions involved in cognitive control of negative emotion,<sup>37</sup> including the dorsolateral prefrontal cortex (Brodmann areas 9 and 46) and ventromedial prefrontal cortex (10 mm radius spherical ROI centered at  $[x,y,z]=[4,32,-5]$ ).<sup>38</sup> Behavioral treatment effects and initial responder/nonresponder differences on demographic and clinical data were calculated using the STATA statistical software, v. 13.1 (Stata, College Station, TX, USA).

## Support vector machine learning

SVM analyses were carried out using the Pattern Recognition for Neuroimaging Toolbox.<sup>39</sup> For each ROI or network, pretreatment contrast images of self-referential criticism were summarized using a dot-product kernel matrix, which was then used as input to a linear SVM. Each SVM was embedded within a leave-one-participant-out cross-validation framework and for each cross-validation fold, data were centered using the mean of the training data. The SVM soft margin parameter *C* was fixed to its default value 1.<sup>39</sup> SVMs were trained to separate CGI-I responders from nonresponders at 1-year follow-up. We employed SVMs because they provide excellent performance for neuroimaging data relative to alternative approaches such as random forests.<sup>40</sup> We tested the statistical significance of each classifier using permutation testing. This involved repeating the cross-validation procedure (1000 times) after randomly permuting the class labels. Significance of the classification accuracy was then derived by counting the number of times the permuted accuracy exceeded the true accuracy and dividing by 1000.

We used a discriminative mapping approach to visualize the relative contribution of the different brain regions to the classifier decision. This is a standard approach in SVM studies using neuroimaging data<sup>39</sup> and was achieved by mapping the predictive weights in the original voxel space. It is important to emphasize that the voxel weights should not be interpreted as describing focal differences between classes in the same way as in mass-univariate analysis. Instead, the weights should be interpreted as a multivariate pattern, where the weight indicates the direction and quantifies the contribution of each voxel to the classifier decision. They do not permit inference at the level of individual voxels.

As a complement to SVM analyses on the dichotomous CGI-I measure, we also conducted relevance vector regressions<sup>41</sup> in Pattern Recognition for Neuroimaging Toolbox with the aim to predict improvement (pre follow-up) on the continuous LSAS-SR.

In the main analyses we found the ACC to be highly predictive of long-term response. Because the ACC has strong anatomical connection with the amygdala,<sup>42</sup> and because it is a functionally heterogeneous region, for example, with regard to control and expression of negative emotion,<sup>43</sup> we analyzed the dACC and vACC subregions separately, and together with the amygdala. Furthermore, follow-up analyses (guided by the significant SVM classifications) using mass-univariate voxel-wise comparisons of initial BOLD response to self-referential criticism were performed between responders and nonresponders, and psychophysiological interactions<sup>44</sup> were conducted using the dACC and vACC as seeds to evaluate couplings between these regions and the amygdala during self-referential criticism.

## Supplementary analyses

Additional SVM analyses were performed using an intention-to-treat approach, that is, all participants that withdrew from the study were



deemed as nonresponders. We also excluded all participants with concurrent psychotropic medication. In addition, ROIs from two previously reported univariate prediction studies of short-term outcome of CBT for SAD<sup>9,10</sup> were evaluated, as well as whole-brain analysis. SVM prediction of SAD diagnostic status, assessed with the structured clinical interview for DSM-IV,<sup>29</sup> was also performed. In addition, we checked for associations between descriptive variables (Table 1) and responder status, withdrawal and treatment compliance. Furthermore, the predictive value of descriptive (clinical and demographic) variables on long-term clinical response was assessed by multiple logistic regression analyses. Results of these analyses are reported in Supplementary Material.

## RESULTS

### Prediction of long-term clinical response

There were 12/23 (52%) responders at 1-year follow-up according to the clinical interview CGI-I. Participants improved on the LSAS-SR from pretreatment to follow-up ( $t(23)=7.52$ ,  $P < 0.001$ , mean  $75.37 \pm 19.1$  to  $44.67 \pm 22.8$ ), and there was no significant change between the posttreatment and 1-year follow-up assessments ( $t(23) = 0.70$ ,  $P = 0.489$ ).<sup>25</sup>

SVM analyses of pretreatment BOLD response to self-referential criticism (self > other referred criticism) showed that information from the ACC was highly accurate in classifying CGI-I responder status 1 year after treatment (balanced accuracy 91.7%, confidence interval 95% 73.2–97.6; Table 2 and Figure 1). Relevance vector regression predictions of improvement on the LSAS-SR were not significant ( $P > 0.100$ ).

SVM analyses taking ACC subdivisions into account showed that the dACC alone (balanced accuracy 86.7%, confidence interval 95% 67.9–95.5,  $P = 0.001$ , area under the receiver-operating characteristic curve (AUC) = 0.97) and dACC together with the amygdala (balanced accuracy 91.7%, confidence interval 95% 73.2–97.6,  $P = 0.001$ , AUC = 0.89), but not the vACC, predicted 1-year treatment response. Furthermore, psychophysiological interaction analyses showed that the dACC, but not the vACC, was significantly (family-wise error, FWE corrected) less coupled with the right amygdala during self-referential criticism at pretreatment ( $Z = 2.91$ ,  $P^{\text{FWE}} = 0.036$ ,  $729 \text{ mm}^3$ ,  $x,y,z$  [30, -1, -14], see also Figure 2) in long-term responders relative to nonresponders, with a similar tendency for dACC–left amygdala coupling ( $Z = 1.81$ ,  $P^{\text{uncorrected}} = 0.035$ ,  $54 \text{ mm}^3$ ,  $x,y,z$  [-21, -1, -11]). Also, pretreatment voxel-wise univariate analyses suggested that the dACC was less reactive to self-referential criticism in responders as compared with nonresponders, although this was significant only when participants on concurrent medication were excluded from the analysis ( $Z = 3.48$ ,  $P^{\text{FWE}} = 0.030$ ,  $4320 \text{ mm}^3$ ,  $x,y,z$  [12, 32, 28]).

**Table 2.** Predictions of clinical outcome at 1-year follow-up. The sensitivity, specificity and balanced classification accuracy (arithmetic mean of sensitivity and specificity) are presented as percentages

	Balanced accuracy	P (balanced) <sup>a</sup>	Sensitivity	Specificity	AUC
<b>ACC</b>	<b>91.7</b>	<b>0.001</b>	<b>83.3</b>	<b>100.0</b>	<b>0.91</b>
Amygdala	47.7	0.531	50.0	45.5	0.46
dIPFC	43.2	0.638	50.0	36.4	0.46
Hippocampus	51.9	0.412	58.3	45.5	0.37
Insula	43.6	0.592	41.7	45.5	0.45
vmPFC	39.0	0.694	41.7	36.4	0.29

Abbreviations: ACC, anterior cingulate cortex; AUC, area under the receiver-operating characteristic curve; dIPFC, dorsolateral prefrontal cortex; vmPFC, ventromedial prefrontal cortex. <sup>a</sup>P-values are calculated from permutation testing with 1000 permutations. Significant balanced accuracies are in bold.

### Supplementary analyses

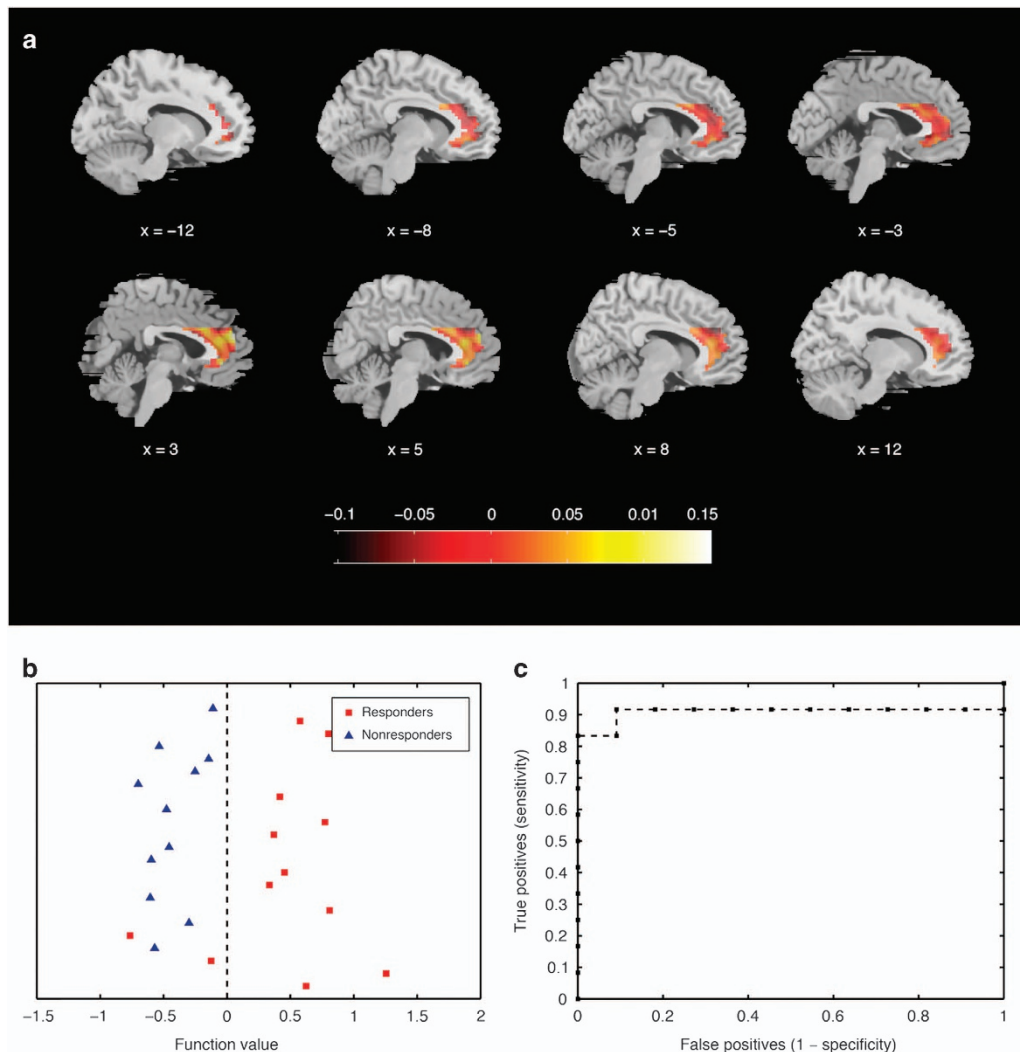
Consistent with the main results, supplementary SVM analyses using the intention-to-treat approach revealed accurate prediction of 1-year iCBT responder status based on information from the ACC, and this region also remained a significant predictor when participants on concurrent psychotropic medication were excluded (Supplementary Material). Neither whole-brain nor ROI analyses using brain regions implicated in previous univariate SAD-fMRI prediction studies of CBT outcome<sup>9,10</sup> were predictive of long-term response, and diagnostic status, according to the structured clinical interview for DSM-IV, could not be predicted. Descriptive characteristics in Table 1 and scanner movement parameters did not differ between responders and nonresponders at pretreatment. Also, multiple logistic regressions using clinical and demographic variables did not reveal significant associations with long-term response to iCBT (see Supplementary Material).

## DISCUSSION

By use of fMRI and SVM classification, we demonstrate that initial multivariate patterns of BOLD response to self-referential criticism in the ACC accurately predicted long-term response to iCBT. Further analyses showed that particularly the dorsal part of the ACC, together with the amygdala, predicted 1-year responder status, and that the initial coupling between these regions, as measured by psychophysiological interactions, was significantly lower in responders relative to nonresponders.

Thus, in individuals with SAD, treatment response 1 year after iCBT was accurately predicted by initial multi-voxel patterns of BOLD response to self-referential criticism in the ACC (92% balanced accuracy, defined as the mean of sensitivity and specificity). Our results are consistent with the notion that ACC-dependent self-referential processing is relevant for treatment outcome.<sup>45</sup> Our long-term ACC findings are also coherent with prior studies using mass-univariate approaches or SVM to predict short-term response to CBT for SAD,<sup>10</sup> panic disorder<sup>17</sup> and depression.<sup>18,19,45</sup> The ACC is, however, a functionally heterogeneous region linked to diverse processes such as conflict monitoring,<sup>46</sup> action-outcome evaluation,<sup>47</sup> fear expression<sup>48</sup> and emotion regulation.<sup>43</sup> Future studies are needed to determine whether such processes are involved in the current experimental task.<sup>27</sup> Meanwhile, refined analyses in this study revealed that the dACC, in contrast to the vACC, was highly predictive, both by itself and together with the amygdala (92% balanced accuracy). Also, voxel-wise univariate analyses suggested lower initial dACC–amygdala coupling, and arguably also lower initial dACC reactivity to self-referential criticism, in responders relative to nonresponders. Both the dACC and the amygdala have been associated with fear expression.<sup>43,48,49</sup> For instance, Vogt *et al.*<sup>50</sup> examined the whole cingulate cortex and showed that 60% of the fear-induced activity was located in the dorsal ACC subregion, suggesting that multivariate brain activation patterns particularly in fear-expressing regions may influence sustained success of CBT.

Prior studies using clinical predictors of treatment response, such as symptom severity, have reported mixed results in SAD.<sup>51</sup> The supplementary analyses in the present study showed that clinical and demographic variables, in contrast to SVM-fMRI, failed to predict iCBT long-term outcome, consistent with the notion that neuroimaging biomarkers add substantially to the predictive value of conventional factors for clinical treatment response.<sup>9</sup> This was also noted in an fMRI random forest study on CBT outcome predictors in generalized anxiety disorder and panic disorder.<sup>52</sup> The clinical utility of the multivariate SVM method may be substantial as it can predict accurately at the individual level. An important feature is that this method provides approximately unbiased estimates of generalizability to new cases. This was achieved by training the models on one sample and testing them



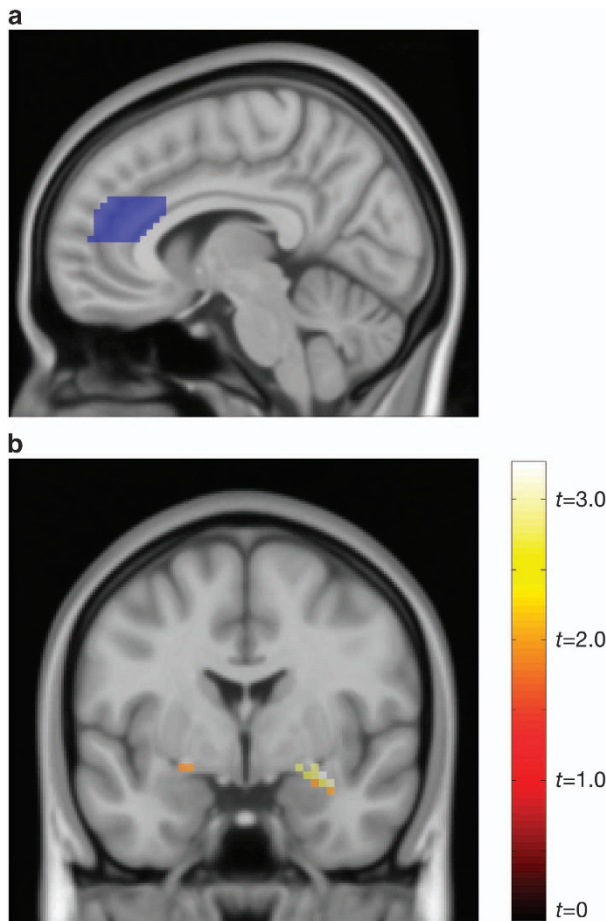
**Figure 1.** Support vector machine classification of responder status at 1-year follow-up in the anterior cingulate cortex. **(a)** Weight map indicating relative weights ascribed to voxels at representative sagittal slices. **(b)** Classification of responder status. **(c)** Receiver-operating characteristic curve, including area under the curve (AUC = 0.91).

on another independent sample, here realized using leave-one-out cross-validation. This represents a first step toward personalized treatment where therapeutic choice is guided by neuroimaging biomarkers. An important next step toward this objective is to validate the SVM models derived here using larger data sets, preferably derived from different sites and/or scanners.

There are several limitations of this study. Because of the small sample size, the results should be interpreted with some caution even though our sample size is in the higher range in comparison with other neuroimaging treatment prediction studies in the anxiety disorders.<sup>20</sup> Unfortunately, reliability assessment of the CGI-I was not performed although the long-term clinical outcome is comparable to previous large-scale randomized controlled trials of iCBT for SAD (including 1-year follow-ups).<sup>5,28</sup> In contrast to the CGI-I, we were not able to predict improvement on the LSAS-SR continuous outcome and prediction of SAD diagnostic status, using the structured clinical interview for DSM-IV, was also not significant, suggesting that treatment outcome measures should be carefully selected in SVM prediction studies. The dACC-amygdala activity may be predictive of change in global symptoms rather than specific social anxiety symptoms, although other reasons for the discrepancy should be noted, such as the use of categorical vs continuous measures, and SVM vs relevance vector regression machine-learning algorithms. Only 13

participants were treatment-naïve at the first fMRI assessment, and possible drug  $\times$  iCBT interactions cannot be ruled out entirely. However, the distribution of participants on concurrent psychotropic medication was not different between responders and nonresponders and, importantly, the accurate predictions from the ACC remained significant when participants on concurrent SSRI treatment were excluded from the main SVM analyses (see Supplementary Material). Also, the supplementary intention-to-treat approach yielded consistent significant predictions by the ACC. The generalizability of the present study, however, may be limited because the SVM classification algorithm has not been applied to a separate clinical sample and our results may also be specific for iCBT and/or the specific fMRI task. Thus, future SVM research should investigate the predictive power of different experimental tasks, including resting-state fMRI, on CBT delivered both via the Internet and face-to-face.

In conclusion, we demonstrate that the initial dACC-amygdala BOLD response pattern to self-referential criticism is predictive of long-term treatment response to iCBT, and that lower pretreatment coupling between these fear-expressing brain regions is associated with stable symptom improvement. In contrast, clinical or demographic variables were not associated with sustained treatment response, underscoring the importance of taking into account neuroimaging biomarkers in prediction studies. To our



**Figure 2.** Statistical parametric map depicting less coupling between the dorsal anterior cingulate cortex (seed region) and the amygdala during self-directed criticism in responders as compared with nonresponders. **(a)** Sagittal view demonstrating the dorsal anterior cingulate cortex (dACC) mask used as the seed region in the psychophysiological interaction (PPI) analysis. **(b)** Coronal view showing amygdala task-dependent coupling with the dACC.

knowledge, this is the first study using the multivariate SVM-fMRI method to successfully predict long-term treatment response in an anxiety disorder. Although SVMs hold high promise, additional studies involving the integration of imaging and related data from more than one site will allow us to further evaluate the clinical utility of this approach.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGMENTS

This study was supported by grants from the Swedish Research Council (Dr Furmark), Linköping University (Dr Andersson), Swedish Research Council for Health, Working Life and Welfare (Dr Carlbring and Dr Furmark) and PRIMA Psychiatry Research Foundation (MSc Månsson). Dr Marquand gratefully acknowledges support from King's College London Centre of Excellence in Medical Engineering, funded by the Wellcome Trust and the Engineering and Physical Sciences Research Council under Grant number WT088641/Z/09/Z and also the NWO under the Language in Interaction project. We are grateful to the staff of the Umeå Functional Brain Imaging Centre and AnnKatrine Larsson for providing excellent research conditions. Dr Mats Fredrikson for valuable input on the manuscript, and Dr Karina Blair for providing word-stimuli to the experimental task.

#### DISCLAIMER

The funding organizations had no role in the study setup, data interpretation or reporting.

#### REFERENCES

- 1 Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry* 2005; **62**: 593–602.
- 2 Beesdo K, Bittner A, Pine DS, Stein MB, Höfler M, Lieb R *et al*. Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Arch Gen Psychiatry* 2007; **64**: 903–912.
- 3 Clark DM Wells A. A cognitive model of social phobia. Heimberg, RG, Liebowitz, MR, Hope, DA, Schneier, FR (eds). *Social phobia: Diagnosis, Assessment and Treatment*. New York, NY, USA: Guilford Press, 1995, pp 69–93.
- 4 Clark DM, Ehlers A, McManus F, Hackmann A, Fennell M, Campbell H *et al*. Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. *J Consult Clin Psychol* 2003; **71**: 1058–1067.
- 5 Andersson G, Carlbring P, Holmström A, Sparthar E, Furmark T, Nilsson-Ihrfelt E *et al*. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. *J Consult Clin Psychol* 2006; **74**: 677–686.
- 6 Hedman E, Furmark T, Carlbring P, Ljótsson B, Rück C, Lindfors N *et al*. A 5-Year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. *J Med Internet Res* 2011; **13**: e39–e39.
- 7 Hedman E, Andersson G, Ljótsson B, Rück C, Mörtberg E, Lindfors N. Internet-based cognitive behavior therapy vs cognitive behavioral group therapy for social anxiety disorder: a randomized controlled non-inferiority trial. *PLoS One* 2011; **6**: e18001.
- 8 Ponniah K, Hollon SD. Empirically supported psychological interventions for social phobia in adults: a qualitative review of randomized controlled trials. *Psychol Med* 2007; **38**: 3–14.
- 9 Doehrmann O, Ghosh S, Polli F, Reynolds G, Horn F, Keshavan A *et al*. Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry* 2012; **70**: 87–97.
- 10 Klumpp H, Fitzgerald DA, Phan KL. Neural predictors and mechanisms of cognitive behavioral therapy on threat processing in social anxiety disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 2013; **45**: 83–91.
- 11 Furmark T, Tillfors M, Marteinsdottir I, Fischer H, Pissiota A, Långström B *et al*. Common changes in cerebral blood flow in patients with social phobia treated with citalopram or cognitive-behavioral therapy. *Arch Gen Psychiatry* 2002; **59**: 425–433.
- 12 Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev* 2012; **36**: 1140–1152.
- 13 Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; **20**: 273–297.
- 14 Pantazatos SP, Talati A, Schneider FR, Hirsch J. Reduced anterior temporal and hippocampal functional connectivity during face processing discriminates individuals with social anxiety disorder from healthy controls and panic disorder, and increases following treatment. *Neuropsychopharmacol* 2014; **39**: 425–434.
- 15 Frick A, Gingnell M, Marquand AF, Howner K, Fischer H, Kristiansson M *et al*. Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. *Behav Brain Res* 2014; **259**: 330–335.
- 16 Shin LM, Liberzon I. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacol* 2009; **35**: 169–191.
- 17 Lueken U, Straube B, Konrad C, Wittchen H-U, Ströhle A, Wittmann A *et al*. Neural substrates of treatment response to cognitive-behavioral therapy in panic disorder with agoraphobia. *Am J Psychiatry* 2013; **170**: 1345–1355.
- 18 Siegle G, Carter C, Thase M. Use of fMRI to predict recovery from unipolar depression with cognitive behavior therapy. *Am J Psychiatry* 2006; **163**: 735–738.
- 19 Fu CHY, Steiner H, Costafreda SG. Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. *Neurobiol Dis* 2013; **52**: 75–83.
- 20 Shin LM, Davis FC, VanElzakker MB, Dahlgren MK, Dubois SJ. Neuroimaging predictors of treatment response in anxiety disorders. *Biol Mood Anxiety Disord* 2013; **3**: 15.
- 21 Hartley CA, Phelps EA. Changing fear: The neurocircuitry of emotion regulation. *Neuropsychopharmacol* 2009; **35**: 136–146.
- 22 Goldin PR. Impact of cognitive behavioral therapy for social anxiety disorder on the neural dynamics of cognitive reappraisal of negative self-beliefs. *JAMA Psychiatry* 2013; **70**: 1048–1056.
- 23 Gong Q, Wu Q, Scarpazza C, Lui S, Jia Z, Marquand A *et al*. Prognostic prediction of therapeutic response in depression using high-field MR imaging. *NeuroImage* 2011; **55**: 1497–1503.



- 24 Steinert C, Hofmann M, Leichsenring F, Kruse J. What do we know today about the prospective long-term course of social anxiety disorder? A systematic literature review. *J Anxiety Disord* 2013; **27**: 692–702.
- 25 Månsson KNT, Carlbring P, Frick A, Engman J, Olsson C-J, Bodlund O *et al*. Altered neural correlates of affective processing after internet-delivered cognitive behavior therapy for social anxiety disorder. *Psychiatry Res* 2013; **214**: 229–237.
- 26 Boettcher J, Hasselrot J, Sund E, Andersson G, Carlbring P. Combining attention training with internet-based cognitive-behavioural self-help for social anxiety: a randomised controlled trial. *Cogn Behav Ther* 2014; **43**: 34–48.
- 27 Blair K, Geraci M, Devido J, McCaffrey D, Chen G, Vythilingam M *et al*. Neural response to self- and other referential praise and criticism in generalized social phobia. *Arch Gen Psychiatry* 2008; **65**: 1176–1184.
- 28 Furmark T, Carlbring P, Hedman E, Sonnenstein A, Clevberger P, Bohman B *et al*. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *Br J Psychiatry* 2009; **195**: 440–447.
- 29 First M, Gibbon M, Spitzer R, Williams J. *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)*. American Psychiatric Press: Washington DC, USA, 1997.
- 30 Zaidar TI, Heimberg RG, Fresco DM, Schneier FR, Liebowitz MR. Evaluation of the clinical global impression scale among individuals with social anxiety disorder. *Psychol Med* 2003; **33**: 611–622.
- 31 Liebowitz MR. Social phobia. *Mod Probl Pharmacopsychiatr* 1987; **22**: 141–173.
- 32 Andersson G, Carlbring P, Furmark T Group SOFIER. Therapist experience and knowledge acquisition in internet-delivered CBT for social anxiety disorder: a randomized controlled trial. *PLoS One* 2012; **7**: e37411.
- 33 Carlbring P, Gunnarsdóttir M, Hedensjö L, Andersson G, Ekselius L, Furmark T. Treatment of social phobia: randomised trial of internet-delivered cognitive-behavioural therapy with telephone support. *Br J Psychiatry* 2007; **190**: 123–128.
- 34 MacLeod C, Mathews A. Cognitive bias modification approaches to anxiety. *Annu Rev Clin Psychol* 2012; **8**: 189–217.
- 35 Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* 2003; **19**: 1233–1239.
- 36 Beckmann M, Johansen-Berg H, Rushworth MFS. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci* 2009; **29**: 1175–1190.
- 37 Ochsner KN, Silvers JA, Huhle JT. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann N Y Acad Sci* 2012; **1251**: E1–24.
- 38 Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol Psychiatry* 2007; **62**: 446–454.
- 39 Schrouff J, Rosa MJ, Rondina JM, Marquand AF, Chu C, Ashburner J *et al*. PRoNTO: pattern recognition for neuroimaging toolbox. *Neuroinformatics* 2013; **11**: 319–337.
- 40 Khondoker M, Dobson R, Skirrow C, Simmons A, Stahl D. A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. *Stat Methods Med Res* 2013. e-pub ahead of print.
- 41 Tipping ME. Sparse Bayesian learning and the relevance vector machine. *JMLR* 2001; **1**: 211–244.
- 42 Ghashghaei HT, Hilgetag CC, Barbas H. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *NeuroImage* 2007; **34**: 905–923.
- 43 Etkin A, Egner T, Kalisch R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci* 2011; **15**: 85–93.
- 44 Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage* 1997; **6**: 218–229.
- 45 Pizzagalli DA. Frontocingulate dysfunction in depression: toward biomarkers of treatment response. *Neuropsychopharmacol* 2011; **36**: 183–206.
- 46 Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* 2004; **8**: 539–546.
- 47 Bush G, Vogt BA, Holmes J, Dale AM, Greve D, Jenike MA *et al*. Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proc Natl Acad Sci USA* 2002; **99**: 523–528.
- 48 Milad MR, Quirk GJ, Pitman RK, Orr SP, Fischl B, Rauch SL. A role for the human dorsal anterior cingulate cortex in fear expression. *Biol Psychiatry* 2007; **62**: 1191–1194.
- 49 LeDoux J. The emotional brain, fear, and the amygdala. *Cell Mol Neurobiol* 2003; **23**: 727–738.
- 50 Vogt BA, Berger GR, Derbyshire SWG. Structural and functional dichotomy of human midcingulate cortex. *Eur J Neurosci* 2003; **18**: 3134–3144.
- 51 Hedman E, Andersson E, Ljótsson B, Andersson G, Andersson E, Schalling M *et al*. Clinical and genetic outcome determinants of internet- and group-based cognitive behavior therapy for social anxiety disorder. *Acta Psychiatr Scand* 2012; **126**: 126–136.
- 52 Ball TM, Stein MB, Ramsawh HJ, Campbell-Sills L, Paulus MP. Single-subject anxiety treatment outcome prediction using functional neuroimaging. *Neuropsychopharmacology* 2014; **39**: 1254–1261.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)