Attentional Bias Modification Alters fMRI Response towards Negative Stimuli in Residual Depression

E. Hilland[1,2], N.I. Landrø[1,2], C. J. Harmer[1,3], M. Browning[4,5], L. A. Maglanoc[1,6], & R. Jonassen[1]

[1]Clinical Neuroscience Research Group, Department of Psychology, University of Oslo, Norway. [2] Diakonhjemmet Hospital, Division of Psychiatry, Oslo, Norway. [3]Psychopharmacology and Emotional Research Lab (PERL), University Department of Psychiatry, University of Oxford, UK. [4]Computational psychiatry laboratory, Department of Psychiatry, University of Oxford. [5]Oxford health NHS Trust, Oxford. [6]NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Norway.

Corresponding authors: Eva Hilland & Nils Inge Landrø
e-mail: eva.hilland@psykologi.uio.no; n.i.landro@psykologi.uio.no
University of Oslo, Department of Psychology
P.O Box 1094
Blindern
0317 Oslo, Norway

Abstract

Background
Modification of attentional biases (ABM) in residual depression may lead to more adaptive emotion perception and emotion regulation reflected in changes in brain activity. There are growing efforts to characterize the neural underpinnings of ABM in depression, but task related emotion processing has so far not been investigated in a larger sample.

Methods
A total of 134 previously depressed individuals with residual symptoms were allocated to 14 days of ABM or placebo in a pre-registered RCT followed by an fMRI modified emotion perception task.

Results
ABM was associated with reduced amygdala and anterior cingulate cortex (ACC) activation compared to placebo during negative images. Response within the insular cortex was associated with the induction of positive affective bias following ABM and with improvement in symptoms.

Conclusions
ABM training has an early effect on brain function within circuitry associated with emotional appraisal and the generation of affective states.

Introduction

A number of effective treatments exist for major depressive disorder. However, following successful treatment relapse is common with 50-70% of patients relapsing within 5 years (1, 2). Residual symptoms are among the strongest predictors for relapse in recurrent depression (3). Cognitive theories of depression posit that biased information processing for emotional stimuli plays a key role in development and relapse in depression (4). Corroborating this, it has been reported that clinically depressed subjects, as well as currently euthymic previously depressed subjects, orient their attention toward negative faces rather than neutral or positive faces (5). Biased attention (AB) and deficits in cognitive control may interfere with emotion regulation and mood state(4). Negative cognitive biases in depression are thought to be facilitated by increased influence from subcortical emotion processing regions combined with attenuated top-down cognitive control.

Computerized Attention Bias Modification (ABM) procedures aim to implicitly retrain biased attentional patterns (6). Although there is debate about the true effect size of ABM in depression (7, 8) some studies have reported reduced depressive symptoms after successful modification of AB (9-12). The neural basis of changes in AB which is believed to be the mechanism of change behind symptom improvement after ABM training, has so far not been investigated. The functional neurobiology of emotion perception distinguishes between structures important for appraisal, generation of affective states and emotion regulation. The amygdala and insular cortex is particularly important within a ventral system linked to the emotional significance of stimuli, and the production of affective states. The ventral ACC plays a main role in automatic regulation of emotional responses. A dorsal system includes the dorsal ACC and prefrontal regions and is argued to be involved in effortful regulation of affective states and subsequent behavior (13, 14).

Single session ABM has previously been linked to lateral prefrontal cortex reactivity towards emotional stimuli in healthy individuals (15) indicating moderation of the dorsal neurocircuitry in emotion perception. One single fMRI study in young women with sub-threshold depression found differences between ABM and placebo in measures of spontaneous fluctuations within the right anterior insula and right middle frontal gyrus (16), areas critical for emotion generation and automatic emotion regulation of emotional responses. Beevers et al. (17) found no neural differences between ABM and placebo in clinical depression as measured by resting state fMRI, but found a pre- to post-training change in resting state connectivity within a network associated with sustained attention to visual information in the placebo group. Overall, these early results provide some evidence that ABM modifies function in emotional regulatory systems although the small study sample sizes and variety of approaches used mean that no consistent effects have been demonstrated.

No study has investigated ABM induced changes in emotion processing using fMRI in a large clinical sample after longer training periods. In this pre-registered clinical trial (NCT02931487) we used a sample of 134 participants previously treated for depression and with various degrees of residual symptoms. Imaging the changes in AB could have a potential translational value since neural correlates of early changes in emotional processing might be a marker of clinical response (18). Given the notion that attention may play a key role in regulating emotional experience (19), we used a classical emotion regulation paradigm to assess aspects of emotion processing after ABM. A main aim was to explore the neural underpinnings of ABM within both ventral- and dorsal emotion perception circuitry. We measured BOLD response to emotionally arousing negative relative to neutral stimuli and when participants attempted to actively regulate their emotional response. Furthermore, we examined how changes in AB, the mechanism by which ABM is believed to work, differs between groups and if potential symptom improvements after ABM would be reflected in brain activity when exposed to negative stimuli.

Methods and materials

Participants and Screening procedures: Patients previously treated for MDD were randomized into two treatment conditions with either a positive ABM- or a closely matched active placebo training. Block randomization was performed at inclusion to ensure equal numbers of participants and similar characteristics for the two groups. Participants were invited to be part of the fMRI study immediately after training and preferably within one week after ABM training. The current clinical trial (NCT02931487) is an extension of a larger double blinded randomized clinical trial (RCT)(NCT02658682) with 322 patients with a history of depression. A total of 136 eligible participants between 18-65 years old were recruited from the main ABM RCT to the MRI RCT.

The main recruitment base was an outpatient clinic in the Department of Psychiatry, Diakonhjemmet Hospital in Oslo. Participants were also recruited from other clinical sites and via social media. Individuals diagnosed with current- or former neurological disorder, psychosis, bipolar spectrum disorders, substance use disorders, attention deficit disorder, or head trauma were excluded via pre-screening. Informed consent was obtained before enrolment. The procedure was approved by The Regional Ethical Committee for Medical and Health Research for Southern Norway (2014/217/REK sør-øst D).

Inclusion criteria were individuals that had experienced more than one depressive episode fulfilling the M.I.N.I A1a (depressed mood) and/or A2a (loss of interest or pleasure) criteria, more than 5 positive on A3 and filling the A5 criterion (DSM 296.30-296.36 Recurrent/ ICD-10 F33.x). To assess both clinically- and subjective evaluations of symptom severity Beck Depression Inventory (BDI-II) (20) and Hamilton Rating Scale for Depression (HDRS) (21) were administered.

Attentional bias modification procedure: The ABM task was a computerized visual dot-probe procedure developed by Browning and coworkers (9). A fixation cross was initially displayed followed by two images (the stimuli) presented concurrently on the top and bottom of the computer screen. Following stimulus onset, a probe (one or two dots) immediately appeared on the same location as one of the image stimuli and remained on the screen until the participant responded. The types of stimuli were pictures of emotional faces of three valences; positive (happy), neutral, or negative (angry and fearful). A single session of the task involved 96 trials with equal numbers of the three stimulus pair types. In addition, equal numbers of trials were randomly presented for 500- or 1000 ms before the probe was displayed. In each trial of the task, stimuli from two valences were displayed, in one of the following pairing types: positive-neutral, positive-negative, and negative-neutral. In the ABM condition, probes were located behind positive stimuli in 87 % of the trials (valid trials), as opposed to 13% with probes located behind the more negative stimuli (invalid trials). Consequently, participants should implicitly learn to deploy their attention toward positive stimuli, and in this way develop a more positive AB when completing the task. The neutral ABM placebo condition was otherwise identical, except the location of the probe, which was located behind the positive (valid trials) stimuli in 50% of the trials. Participants completed two sessions (96 trials) of ABM daily during the course of fourteen days (28 sessions in total) on identical notebook computers (14" HP EliteBook 840, 1600x900, 8GB, Intel Core i5-4310U), which were set up and used exclusively for ABM-training.

MRI Scan acquisition: Scanning was conducted on a 3T Philips Ingenia whole-body scanner, with a 32 channel Philips SENSE head coil (Philips Medical Systems). Functional images were obtained with a single-shot T2* weighted echo planar imaging sequence (repetition time (TR): 2000 ms; slice echo time (TE): 30 ms; field of view (FOV): 240x240x117; imaging matrix: 80x80; flip angle 90°, 39 axial slices, interleaved at 3 mm thickness, no gap, voxel size 3x3x3 mm). The scanning session consisted of 340 volumes, synchronized to the onset of the experiment. Slice orientation was adjusted to be 45° relative to the line running from the anterior to posterior commissure. A T1-weighted anatomical image with a voxel size of 1x1x1 mm was recorded for registration of the functional images (TR: 8.5 ms; TE: 2.3 ms; FOV: 256x256x184; flip angle: 7°;184 sagittal slices).

fMRI Experimental procedure: The study used a modified emotion regulation experiment. Participants were scanned as they were viewing sequences of negative and neutral images while carrying out instructions either to down-regulate their emotional responses using a reappraisal strategy, or to simply allow themselves to attend to the pictures without trying to influence their emotional reactions. After each image the participants provided a rating of the intensity of their emotional state using a visual analogue scale (VAS) ranging from neutral to negative. Stimuli were selected from the International Affective Picture System (22) and the Emotional Picture Set (23). Negative and positive pictures were counterbalanced concerning their valence and arousal scores. Each trial started with a fixation cross followed by a written instruction, ("Attend" or "Regulate"). The instruction was presented for 2000 ms. A negative or neutral image was presented for 6000 ms, followed by a

rating screen time-locked to 6000 ms. Between stimuli there was a temporal jitter randomized from 2000-8000 ms (mean ISI; 3,700 ms) to optimize statistical efficiency in the event related design (24). The task consisted of blocks of 18 trials with a 20 second null-trial between the two blocks. The procedure was completed in two independent runs during the scanning session, 72 trials in total. In each block 12 items were neutral and 24 items were negative, giving three counterbalanced experimental conditions; AttendNeutral, AttendNegative, RegulateNegative. The stimulus-order in each block was interspersed pseudo-randomly from 12 unique lists. The total duration of one single functional scanning run was ~11 minutes, and total scan time~22 minutes. Stimuli were presented using E-Prime 2.0 software (Psychology Software Tools). An MRI compatible monitor for fMRI was placed at the end of the scanner behind the participants' head. Participants watched the monitor using a mirror placed at the head coil. Responses were collected with a response grip with two response buttons. Physiological data (heart and respiration curves) were recorded at 1000 Hz using a clinical monitoring unit digitized together with scanner pulses.

Training and instruction procedures: A written protocol with detailed instructions was used to introduce the emotion regulation experiment. The protocol was dictated for each participant by the researcher outside the MRI-scanner in order to standardize the verbal instructions. The fMRI experiment had three in-scanner exercise trials before scan start in order to make participants familiar with the instructions, timing, response buttons and VAS scale. The training procedure was repeated before the second run of the experiment.

Statistical analysis: Changes in clinically evaluated symptoms were analyzed in PASW 25.0 (IBM) using a univariate ANOVA with intervention (ABM versus placebo). Symptom differences (HRSD) from baseline to follow-up were the dependent variables, and baseline symptom level (using the same measure) was entered as a covariate.

To investigate participants' self-reported emotional reactivity during the experiment (VAS scores) a factor based on the three experimental conditions AttendNeutral, AttendNegative, RegulateNegative was created and analyzed in a repeated measures ANOVA, followed by a post-hoc test. The potential interaction with picture set was explored by adding picture set together with ABM versus Placebo as a second fixed factor.

fMRI data preprocessing and noise reduction: The FMRIB Software Library version (FSL version 6.00) (www.fmrib.ox.ac.uk/fsl) (25, 26) was used to pre-process and analyze fMRI data. FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). In conjunction with FEAT FSL-PNM, 34 EVs were applied to regress out physiological noise from pulse and respiration (27). Registration to high resolution structural and/or standard space images was carried out using FLIRT (28, 29). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (30). All registrations were manually inspected to ensure proper alignment. Time-series statistical analysis was carried out using

FILM with local autocorrelation correction (31). Linear registration was conducted with 12 DOF. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by Z>2.3 and a (corrected) cluster significance threshold of P=0.05 (32). Two participants were excluded from the analyses due to signal loss caused by a technical problem with the head coil. Time series from each subjects' two first level runs were combined using an intermediate fixed effect model in FEAT before submission to second level analysis. A total of 134 participants, 64 from the ABM group and 70 from the placebo group were included in the intermediate - and the higher level FEAT analysis at group level.

Small volume correction (SVC´s): Regions from a recent meta-analysis on neuroimaging and emotion regulation (33) were used to explore differences between ABM and placebo in predefined emotion regulation circuitry (see, 34) . This meta-analysis is comprised of 48 neuroimaging studies of reappraisal where the majority of studies involved downregulation of negative affect. Buhle (33) reported seven clusters related to emotion regulation consistently found within prefrontal cognitive control areas when contrasted to passive viewing of negative images. The clusters were situated in right and left middle frontal gyrus, right inferior frontal gyrus, right medial frontal gyrus, right and left superior temporal lobe, and left middle temporal gyrus. The bilateral amygdala, but no other brain regions was reported for the opposite contrast comparing negative viewing to emotion regulation. Binary spheres with a 5 mm radius based on MNI coordinates of peak voxels were used for the predefined regions. Two single masks were created for each contrast (RegulateNegative > AttendNegative, AttendNegative > RegulateNegative) including the seven cortical spheres and the two subcortical spheres respectively to adjust for multiple comparisons within each contrast condition. Z- threshold was set to 2.3 and cluster p-threshold was .05.

Whole brain analysis: The following contrasts of interests were set up: AttendNegative > AttendNeutral tests whether ABM influences overall activity in response to the presentation of emotional stimuli, AttendNegative > RegulateNegative investigates emotional reactivity compared to regulation strategies, and RegulateNegative > AttendNegative to investigates regulatory processes compared to viewing of negative stimuli. Parameter estimates with baseline symptoms (HRSD) as covariate for all contrast were calculated. In order to test whether ABM influenced the neural systems producing attentional bias and symptom improvement, group differences were further explored in an interaction model in relation to degree of AB change and degree of symptom change.

Spatial smoothing FWHM was set to 5 mm. Featquery was used for FEAT result interrogation. Mean local percent signal change was extracted to explore individual distribution within significant clusters from FEAT, and effect sizes were thereafter estimated by the use of univariate ANOVAs.

Results

Sample characteristics

	Placebo (n=71)	ABM (n=63)	Value	Sig.
Age	39,2 (13.4)	38,6 (12.8)	.077	.782
Gender (females)	46	55	3.068	.080
Education Level (ISCED)	5,92 (1,2)	5,86 (1,2)	.081	.776
Medication (SSRI)	25	23	.087	.768
Number of previous MDE	4,4 (5,1)	4,6 (7,1)	.048	.827
Days between ABM and fMRI	6.9 (8.7)	6.6 (7.2)	.041	.840
Baseline symptoms:				
HRSD	7,63 (4,76)	9,36 (6,15)	3.592	060
BDI-II	12,09 (8,66)	17,40 (11,50)	9.150	.003 **

Table 1 shows means, standard deviations together with F-values and Chi-square tests respectively. MDE=Major Depressive Episodes according to M.I.N.I. SSRI= any current usage of an antidepressant belonging to the Selective Serotonin Reuptake Inhibitors. ISCED= International Standard Classification of Education. P-values from Pearson Chi-Square test is presented for dichotomous variables.

Analyses of symptom change after ABM was corrected for the observed difference in BDI-II. All fMRI analysis was conducted with baseline symptoms (HRSD) as a covariate and the observed BDI-II differences between the ABM group and placebo was explored by post-hoc analysis by exclusion of potential outliers (see Supplemental).

Symptom severity and symptom change after ABM: There was a statistically significant effect of the intervention for rater-evaluated depression as measured by the change in HRSD, with symptoms of depression reducing to a greater extent in the ABM group. This is in accordance with the results reported in the study by Jonassen et al. (35). Mean changes in ABM (M =1.41 (5.60)) and placebo (M= -.13 (6,33)) [$F(1,133) = 4.333$, $\eta^2 = .03$, $p = .032$]. No statistically significant effect was found for self-reported symptoms as measured by the BDI-II (M =3.93 (7.25)) and placebo (M= 2.44 (5.60)) [$F(1,133) = .285$, $p = .594$]. There was general symptom improvement in both ABM and placebo as measured by the BDI-II from baseline (14.6(10.41)) to post training (11.4(10.58)) [$F(1,133) = 31.482$, $\eta^2 = .19$, $p < .001$].

Subjective ratings of negativity: There was a statistically significant difference between self-reported emotional reactivity measured by VAS scores during the fMRI experiment between task conditions. The repeated measures

ANOVA showed that mean VAS scores were lowest when viewing neutral images (M=8.2(7.8)) followed by when patients were encouraged to regulate negative experience towards negative images (M=40.8 (16.9)), and highest for the passive viewing of negative images (62.0(15.3)) [F= (1,133) = .074, η^2= .93, p < .001], but did not differ between ABM and placebo [F= (1,133) = .993, p < .646]. A post hoc test showed that the differences between the passive and regulate viewing conditions for negative stimuli was large and statistically significant F (1,133) = 202.81, η^2= .60, p < .001].

Effects of ABM during conscious emotion regulation: SVC analyses within emotion regulation circuitry revealed more activation in right and left middle frontal gyrus, two of the three largest cortical clusters reported in the meta-analysis (33) in the RegulateNegative>AttendNegative contrast (MNI coordinates xyz = -, 36, -10; Z=3.86; p<.001). For the AttendNegative>RegulateNegative condition there was increased bilateral amygdala activation in both the ABM and placebo group (Supplemental Figure 1), but no difference between ABM and placebo (MNI xyz = -18, -2, -16; voxel size =81; Z=6.71; p=.005; and MNI xyz = 28, 0, 16; voxel size = 57; Z = 4.52; p = .009;).

Effects of ABM in response to negative images: For AttendNegative>AttendNeutral images the placebo group demonstrated greater activation in a cluster within pregenual ACC, the paracingulate – and medial cortex bilaterally, extending to the right frontal orbital cortex and the frontal pole compared to ABM (Figure 1.). The peak activation for the cluster was found in the left pregenual ACC (MNI coordinates xyz =-16, 36, -10; Z=3.86; p=.001).

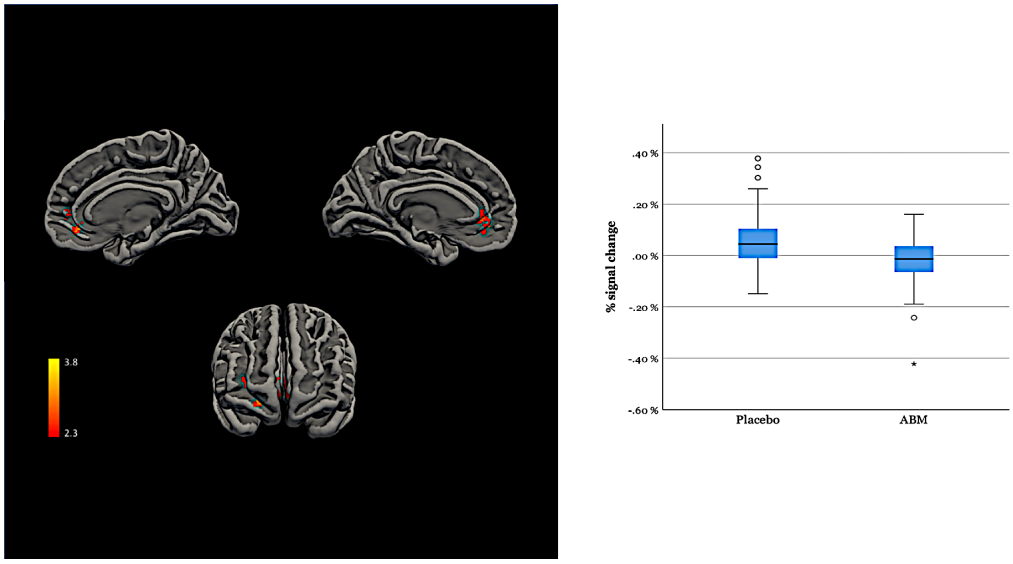


Figure 1 shows cluster activation (Z>2.3) for Placebo over ABM for negative images (left), together with distribution of individual percentage signal changes over significant clusters (right).

SVC corrected for baseline HRSD revealed bilateral amygdala activation for both ABM and placebo, structures linked to the generation of an affective state. The Placebo group had higher right (MNI xyz = -18, -6, -20; Z=2.89;

p=.03) and left (MNI xyz = 28, 0, -16; Z = 2.55; p= .04) amygdala reactivity towards negative images compared to ABM (Figure 2).

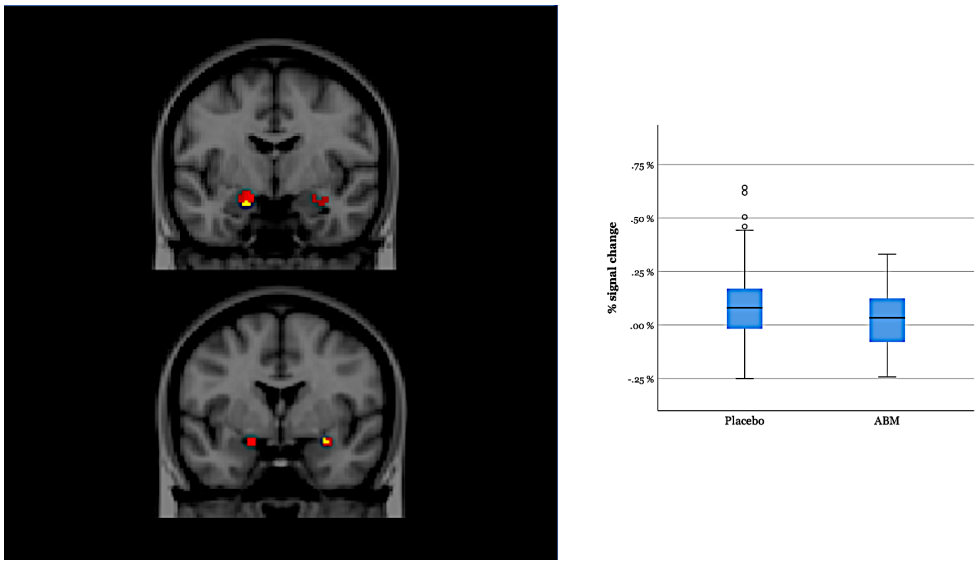


Figure 2 shows amygdala mean activation (left) for placebo and ABM (red) and peak voxels where placebo had more activation compared to ABM for negative images (yellow) and distribution of individual percentage signal changes over significant clusters (right).

Interaction with degree of attentional biases and symptom change: Group differences were explored in an interaction model in relation to degree of AB change and degree of symptom change to test whether ABM influenced the neural systems underlying AB and symptoms. Two distinct clusters were associated with the interaction between the AttendNegative>AttendNeutral condition, the intervention (group) and degree of AB change (MNI xyz = 54,-24, 8; Z=4.05; p<.001) and (MNI xyz =-50, 0 , 10; Z= 3.44; p<.02).

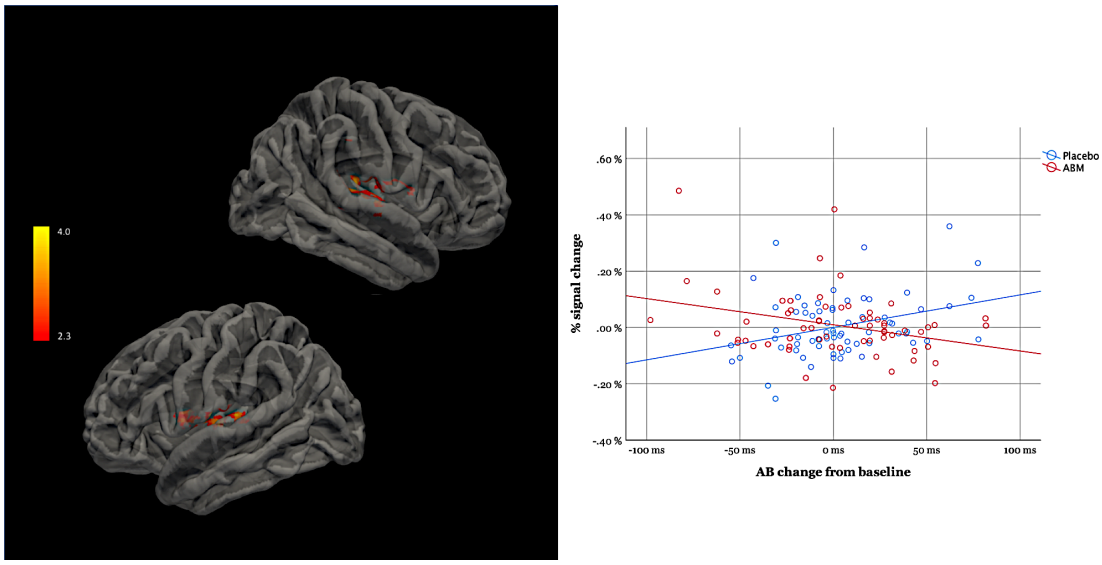


Figure 3 shows areas activated in association to the interaction between AB and the intervention (left). The scatter plot (right) shows the regression lines and individual distribution in the ABM and the placebo condition.

An interaction between negative stimuli, the intervention, and degree of symptom change (SC) were found within the right planum temporale and insular cortex (MNI xyz = 50,-10, 18; Z=5.28; p<.001), also within areas involved in the generation an affective state according to the model by Phillips et al (13) (Figure 4.).

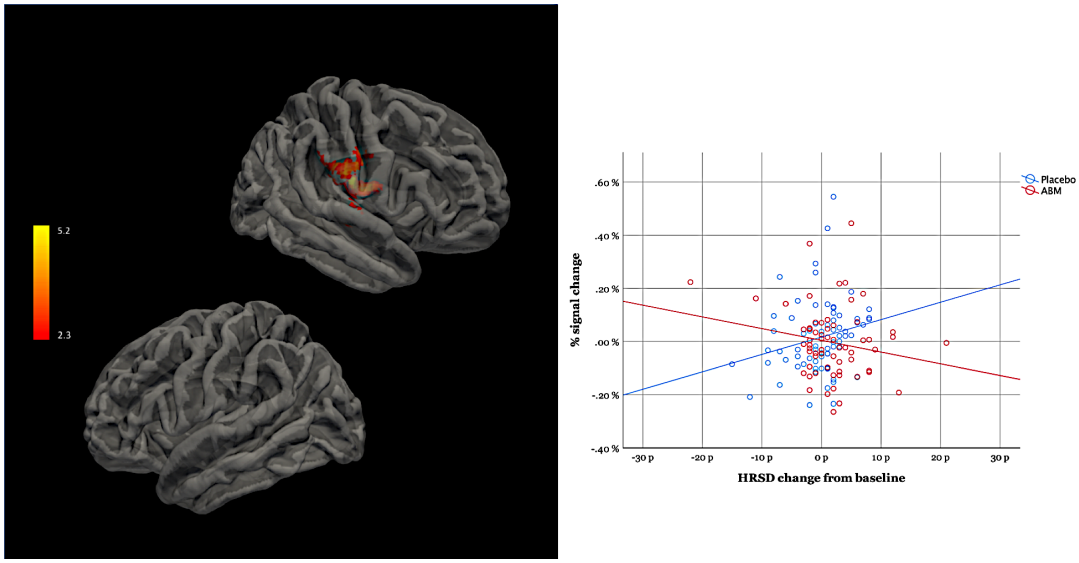


Figure 4 shows areas activated in association to the interaction between HRSD and the intervention (left). The scatter plot (right) shows the regression lines and individual distribution in the ABM and the placebo condition.

Discussion

Our results show attenuated fMRI responses after two weeks of ABM within areas consistently associated with emotional appraisal and the generation of affective states, a circuitry known to be altered in depression (4, 13, 14). The placebo group showed more pronounced activation to negative images in midline structures and pregenual ACC. Analysis of the mechanism of change showed that degree of changes in AB was linearly linked to activity in the insular cortex bilaterally. Symptom improvement after ABM was linearly associated with activation in the right insular cortex. Analyses within predefined areas associated with effortful emotion regulation revealed activation within the left and right middle frontal gyrus across groups (33). The amygdala was also more activated during passive viewing versus active regulation within predefined areas, but did not differ between ABM and placebo. We found no differences between ABM and placebo as measured by subjective ratings of perceived negativity, which may imply that early effects of ABM are restricted to changes in emotion generation and appraisal, as opposed to more conscious forms of emotion regulation linked to the dorsal system. Accordingly, we did find rater evaluated -, but not self-reported changes in symptoms after ABM.

A considerable number of meta-analyses using functional connectivity in depression have shown altered activity in areas that distinguished ABM and

placebo in the current study including the insula and ACC (36-38). In a study by Horn and coworkers (39) increased connectivity between pregenual ACC and insula was found in severely depressed patients compared to mildly depressed patients and healthy controls. Functional connectivity in the insula has been associated with abnormal introceptive activity (40) and fronto-insular connectivity has been linked to maladaptive rumination (41) in depression. Midline brain structures including pregenual ACC has been linked to self-referential processing (42), hopelessness (43), anhedonia (44) and impaired emotion processing (45) in studies of functional connectivity and depression. Notably, the ACC and insula are together with the amygdala core areas of the salience network which determines the significance of external stimuli. The salience network has been hypothesized to play a role in switching between task positive- and negative networks (46, 47) and may well play a role in symptom improvement after ABM as found in this study.

The insula and the amygdala are among core brain areas that respond preferentially to negative stimuli in healthy individuals, and activation in the insula and ACC has repeatedly been reported across a range of experiments using emotional tasks with cognitive demand and mental imagery (48, 49). Neural responses to negative stimuli within the amygdala, insula and ACC are found to be more pronounced in depressed patients versus healthy controls (50). Ma (51) describes an emotional circuit including the insula, bilateral amygdala and ACC affected by antidepressant medication by decreasing activity towards negative- and increasing activity towards positive stimuli. Antidepressants have been hypothesized to work by remediating negative affective biases, i.e. targeting the same mechanism as when applying an ABM procedure (52-54). Similarly, the moderation of awareness towards negative stimuli via ABM (the mechanism of change) may alter automatic emotional vigilance and arousal towards negative stimuli. These moderations may lead to altered parasympathetic responses via circuitry involving the amygdala and ACC. The translation of these changes into improved subjective mood may take some time as the person learns to respond to this new and more positive social and emotional perspective of the world. However, neural correlates of early changes in the processing of emotional stimuli might be a marker of a process leading to symptom improvement. This model is consistent with cognitive theories of depression (4, 55) which the ABM procedure builds on. Accordingly, studies on cognitive behavioral therapy (CBT) shows that pregenual ACC is positively correlated with the degree of symptom improvement (56-61). Moreover, given that that the pregenual ACC is believed to play an important role in downregulation of limbic hyperreactivity (14, 62, 63) the group difference found in this study may reflect more adaptive emotion processing after ABM.

In the current study the treatment groups were well balanced in size, but there was an observed difference in self-reported and clinician rated depression symptoms at baseline. To address the difference we added baseline depression symptoms (HRSD) as covariate for all fMRI analyses (see 64, 65). Additionally, an analysis of symptom change after ABM was corrected for the observed difference in baseline depression symptoms, and a post-hoc exploration was performed to analyze the effect of potential outliers (See Supplemental). The general symptom improvement observed in the placebo

group as well as in the ABM group, suggests that inclusion of a healthy control group together with an assessment only MDD group would have strengthened the potential for casual interpretation in the current study.

Worldwide, there is a pressing demand for evidence-based treatments in mental health. It has been argued that psychotherapy research does not provide explanations for how or why even the most commonly used interventions produce change (66). In a recent statement from the *Lancet Psychiatry's Commission on treatments research in tomorrow's science* the authors argue that there is an acute need to improve treatment and thus clinical trials should focus not only on efficacy, but also on identification of the underlying mechanisms through which treatments operate (67). The current study is addressing such mechanisms by targeting changes in AB which is believed to be the mechanism that translate into symptom improvement after ABM. The current study is based on a large sample of patients, it uses a well validated emotion perception task and follows a stringent pre-registered research protocol which represents a strength. This study exploits the link between a psychological mechanism, clinical measures and underlying brain function measured by fMRI and thus the results should have translational potential.

Conclusion

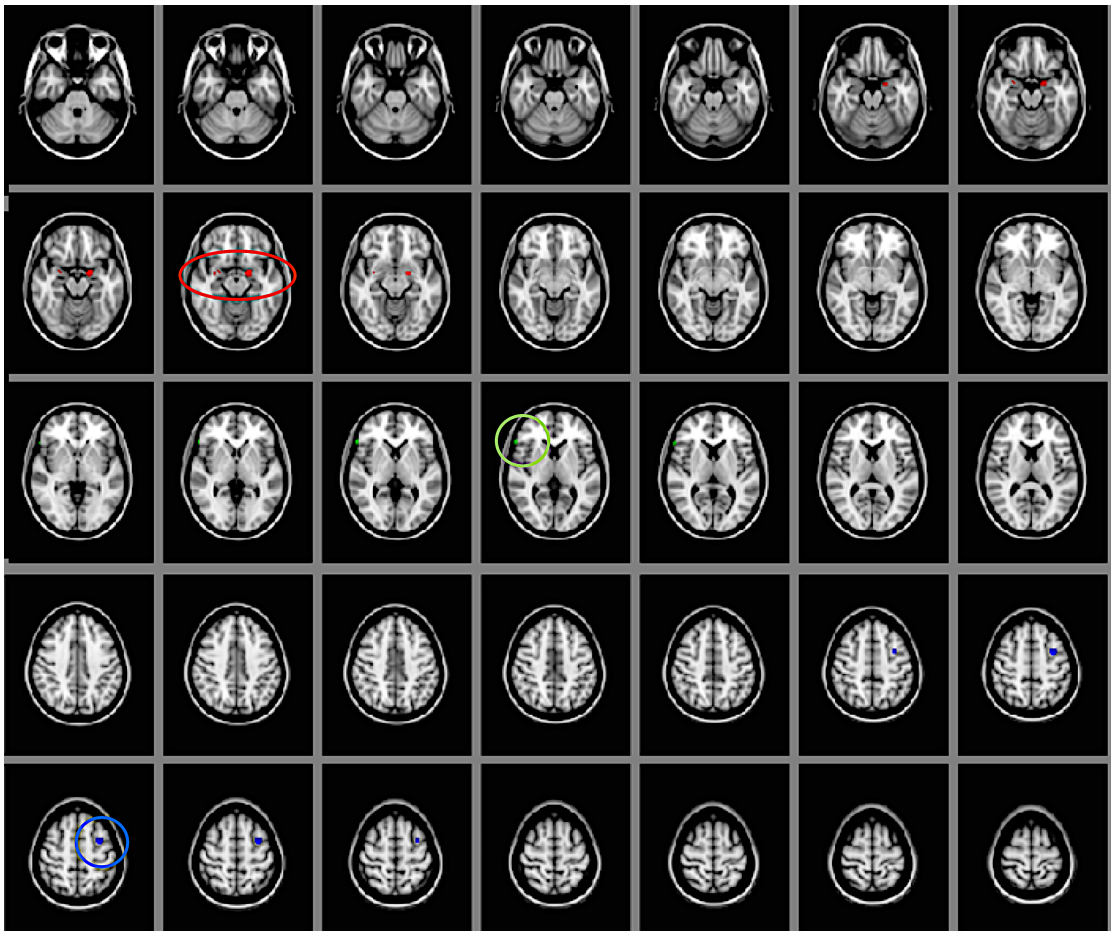
This study demonstrates reduced depressive symptoms following ABM as well as alterations in brain circuitry linked to emotion processing in the context of successful modification of AB. Our findings represent the first experimental evidence bridging the gap between a positive ABM effect and the underlying neural mechanism.

Acknowledgements: We want to thank our fMRI research assistant Dani Beck. Further, we thank Diakonhjemmet Hospital, Division of Psychiatry, for help and support during the recruiting period and the Intervention Center, OUS for radiological assistance in MRI protocols, data acquisitions and screening for unexpected neuropathological findings. We thank Tor Endestad for establishing invaluable infrastructure for MRI research in our department. We also thank our external recruitment sites, Unicare, Coperiosenteret AS, Torgny Syrstad, MD , Synergi Helse AS and Lovisenberg Hospital. The project is supported by the South-Eastern Norway Regional Health Authority, grant number: 2015052, Research Council Norway, grant number: 229135 and Department of Psychology, University of Oslo.

Conflicts of interest: CJH has received consultancy fees form Johnson and Johnson Inc, P1 vital and Lundbeck. MB hold a part time position at P1 vital Ltd and owns shares in P1 vital products Ltd. MB has received travel expenses from Lundbeck and has acted as a consultant for J&J, and NIL has received consultancy fees and travel expenses from Lundbeck.

Supplemental information

Significant clusters from SVC´s



Supplemental Figure 1 shows activation map and cluster index from SVC´s based on the contrasts negative > regulate that revealed bilateral amygdala response (red), and regulate > negative (green and blue).

Supplemental table 1: MNI coordinates ROI analysis

Contrast	Volume (cm3)	*z*	Coordinates x	y	z	Anatomical labels
Rappraise>Look			60	24	3	Middle frontal gyrus R
			51	15	48	Inferior frontal gyrus R
			9	30	39	Medial frontal gyrus R
			-33	3	54	Middle frontal gyrus L
			63	-51	39	Superior parietal lobule R
			-42	-66	42	Superior parietal lobule L
			-51	-39	3	Middle temporal gyrus L
Look>Reapprasie			30	61	28	Amygdala L
			54	61	28	Amygdala R

Clusters of activation for contrasts of interest. We present the volume of the cluster (thresholded at voxelwise $p < 0.01$, uncorrected), the peak voxel (maximum "intensity" z-statistic), MNI coordinates of peak activations within each cluster, Brodman areas (based on local maxima cluster index and Yale BioImage Suite Package http://bisweb.yale.edu/build/mni2tal.html) and the anatomical extent of the cluster (based on local maxima cluster index and Harvard_Oxford Cortical Structural Atlas in FSL). Anatomical labels in bold correspond to the MNI coordinates of peak activations. L, left; R, right

Supplemental table 2: MNI coordinates of activation clusters from whole brain analyses

Contrast	Volume (cm3)	*z*	Coordinates x	y	z	Anatomical labels
Neutral images ABM>Placebo	676	3.88	-2	34	0	**Cingulate gyrus, anterior division (L),** Cingulate Gyrus, anterior division (R), Paracingulate Gyrus (L), Frontal Medial Cortex (L), Frontal Pole (R)
Negative>Neutral Placebo>ABM	1202	3.86	-16	36	-10	**Frontal Medial Cortex (L),** Cingulate Gyrus, anterior division (L,R), Frontal Pole (R), Paracingulate Gyrus (L), Frontal Medial Cortex (L)
Negative>Neutral Intervention x AB	1061	4.05	54	-24	8	**Planum temporale (R),** Insular cortex (L,R), Planum temporale (L), Central opercular cortex (L,R), Heschls Gyrus (L)
Negative>Neutral Intervention x HRSD	872	5.28	50	-10	18	**Central opercular cortex (R),** Insular córtex (R), postcentral gyrus(R), precentral gyrus(R)

Clusters of activation for contrasts of interest. We present the volume of the cluster (thresholded at voxelwise $p < 0.01$, uncorrected), the peak voxel (maximum "intensity" z-statistic), MNI coordinates of peak activations within each cluster, Brodman areas (based on local maxima cluster index and Yale BioImage Suite Package http://bisweb.yale.edu/build/mni2tal.html) and the anatomical extent of the cluster (based on local maxima cluster index and Harvard_Oxford Cortical Structural Atlas in FSL). Anatomical labels in bold correspond to the MNI coordinates of peak activations. L, left; R, right.

Post-hoc assessment of potential outliers

Exclusion of 19 potential outliers that could cause differences in symptom degree and brain activity was conducted using the cut-off for moderate depression (BDI-II > 21), ABM (M =11.1 (6,75)) and placebo (M= 9.7 (5,81)) and did not influence the negative > neutral contrast (Figure 1.) [$F(1,102) = 14.933$, $p < .001$] or the amygdala contrast (Figure 2.) [$F(1,102) = 8.403$, $p = .005$].

Analysis site

fMRI analyses were performed on the Abel super cluster, owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR), and operated by the Department for Research Computing at USIT, the University of Oslo IT-department. http://www.hpc.uio.no/

References

1. Brent DA, Brunwasser SM, Hollon SD, Weersing VR, Clarke GN, Dickerson JF, et al. (2015): Effect of a Cognitive-Behavioral Prevention Program on Depression 6 Years After Implementation Among At-Risk Adolescents: A Randomized Clinical Trial. *JAMA Psychiatry*. 72:1110-1118.
2. Fergusson DM, Boden JM, Horwood LJ (2007): Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes. *Br J Psychiatry*. 191:335-342.
3. Mojtabai R (2001): Residual symptoms and impairment in major depression in the community. *Am J Psychiatry*. 158:1645-1651.
4. Disner SG, Beevers CG, Haigh EA, Beck AT (2011): Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*. 12:467-477.
5. Joormann J, Gotlib IH (2007): Selective attention to emotional faces following recovery from depression. *J Abnorm Psychol*. 116:80-85.
6. MacLeod C, Rutherford E, Campbell L, Ebsworthy G, Holker L (2002): Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of abnormal psychology*. 111:107.
7. Cristea IA, Kok RN, Cuijpers P (2015): Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*. 206:7.
8. Grafton B, MacLeod C, Rudaizky D, Holmes EA, Salemink E, Fox E, et al. (2017): Confusing procedures with process when appraising the impact of cognitive bias modification on emotional vulnerability. *The British Journal of Psychiatry*. 211:266-271.
9. Browning M, Holmes EA, Charles M, Cowen PJ, Harmer CJ (2012): Using attentional bias modification as a cognitive vaccine against depression. *Biological psychiatry*. 72:572-579.
10. Yang W, Ding Z, Dai T, Peng F, Zhang JX (2015): Attention bias modification training in individuals with depressive symptoms: A randomized controlled trial. *Journal of behavior therapy and experimental psychiatry*. 49:101-111.
11. Wells TT, Beevers CG (2010): Biased attention and dysphoria: Manipulating selective attention reduces subsequent depressive symptoms. *Cognition and Emotion*. 24:719-728.
12. Baert S, De Raedt R, Schacht R, Koster EH (2010): Attentional bias training in depression: therapeutic effects depend on depression severity. *Journal of behavior therapy and experimental psychiatry*. 41:265-274.
13. Phillips ML, Drevets WC, Rauch SL, Lane R (2003): Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biol Psychiatry*. 54:515-528.
14. Phillips ML, Ladouceur CD, Drevets WC (2008): A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol Psychiatry*. 13:829, 833-857.

15. Browning M, Holmes EA, Murphy SE, Goodwin GM, Harmer CJ (2010): Lateral prefrontal cortex mediates the cognitive modification of attentional bias. *Biological psychiatry*. 67:919-925.
16. Li H, Wei D, Browning M, Du X, Zhang Q, Qiu J (2015): Attentional bias modification (ABM) training induces spontaneous brain activity changes in young women with subthreshold depression: a randomized controlled trial. *Psychological medicine*.1-12.
17. Beevers CG, Clasen PC, Enock PM, Schnyer DM (2015): Attention bias modification for major depressive disorder: Effects on attention bias, resting state connectivity, and symptom change. *J Abnorm Psychol*. 124:463-475.
18. Godlewska B, Browning M, Norbury R, Cowen PJ, Harmer CJ (2016): Early changes in emotional processing as a marker of clinical response to SSRI treatment in depression. *Translational psychiatry*. 6:e957.
19. Gotlib IH, Joormann J (2010): Cognition and depression: current status and future directions. *Annu Rev Clin Psychol*. 6:285-312.
20. Beck AT, Steer RA, Brown GK (1996): Beck depression inventory-II. *San Antonio*. 78:490-498.
21. Hamilton M (1960): A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 23:56-62.
22. Lang P, Bradley MM, Cuthbert BN (2008): International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL. .
23. Wessa M, Kanske P, Neumeister P, Bode K, Heissler J, Schönfelder S (2010): EmoPics: Subjektive und psychophysiologische Evaluation neuen Bildmaterials für die klinisch-bio-psychologische Forschung. *Zeitschrift für Klinische Psychologie und Psychotherapie*. 39:77.
24. Dale AM (1999): Optimal experimental design for event-related fMRI. *Human brain mapping*. 8:109-114.
25. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. (2004): Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 23:S208-S219.
26. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012): FSL. *Neuroimage*. 62:782-790.
27. Brooks JC, Beckmann CF, Miller KL, Wise RG, Porro CA, Tracey I, et al. (2008): Physiological noise modelling for spinal functional magnetic resonance imaging studies. *Neuroimage*. 39:680-692.
28. Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 17:825-841.
29. Jenkinson M, Smith S (2001): A global optimisation method for robust affine registration of brain images. *Medical image analysis*. 5:143-156.
30. Andersson JL, Jenkinson M, Smith S (2007): Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. *FMRIB Analysis Group of the University of Oxford*. 2.
31. Woolrich MW, Ripley BD, Brady M, Smith SM (2001): Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage*. 14:1370-1386.

32. Worsley K (2001): Statistical analysis of activation images. *Functional MRI: an introduction to methods*. 14:251-270.
33. Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, et al. (2014): Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cerebral Cortex*. 24:2981-2990.
34. Poldrack RA (2007): Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*. 2:67-70.
35. Jonassen R, Harmer CJ, Hilland E, Maglanoc LA, Kraft B, Browning M, et al. (2018): Effects of Attentional Bias Modification on Residual Symptoms in depression. A Randomized Controlled Trial. *bioRxiv (preprint)*.
36. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA (2015): Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. *JAMA psychiatry*. 72:603-611.
37. Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, et al. (2007): Resting-State Functional Connectivity in Major Depression: Abnormally Increased Contributions from Subgenual Cingulate Cortex and Thalamus. *Biological psychiatry*. 62:429-437.
38. Wang L, Hermens D, Hickie I, Lagopoulos J (2012): A systematic review of resting-state functional-MRI studies in major depression. *Journal of affective disorders*. 142:6-12.
39. Horn DI, Yu C, Steiner J, Buchmann J, Kaufmann J, Osoba A, et al. (2010): Glutamatergic and resting-state functional connectivity correlates of severity in major depression–the role of pregenual anterior cingulate cortex and anterior insula. *Frontiers in systems neuroscience*. 4:33.
40. Avery JA, Drevets WC, Moseman SE, Bodurka J, Barcalow JC, Simmons WK (2014): Major depressive disorder is associated with abnormal interoceptive activity and functional connectivity in the insula. *Biological psychiatry*. 76:258-266.
41. Hamilton JP, Furman DJ, Chang C, Thomason ME, Dennis E, Gotlib IH (2011): Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination. *Biological psychiatry*. 70:327-333.
42. Lemogne C, Gorwood P, Bergouignan L, Pélissolo A, Lehéricy S, Fossati P (2010): Negative affectivity, self-referential processing and the cortical midline structures. *Social cognitive and affective neuroscience*. 6:426-433.
43. Grimm S, Ernst J, Boesiger P, Schuepbach D, Hell D, Boeker H, et al. (2009): Increased self‐focus in major depressive disorder is related to neural abnormalities in subcortical‐cortical midline structures. *Human brain mapping*. 30:2617-2627.
44. Walter M, Henning A, Grimm S, Schulte RF, Beck J, Dydak U, et al. (2009): The relationship between aberrant neuronal activation in the pregenual anterior cingulate, altered glutamatergic metabolism, and anhedonia in major depression. *Archives of general psychiatry*. 66:478-486.
45. Grimm S, Boesiger P, Beck J, Schuepbach D, Bermpohl F, Walter M, et al. (2009): Altered negative BOLD responses in the default-mode network during emotion processing in depressed subjects. *Neuropsychopharmacology*. 34:932.

46. Menon V, Uddin LQ (2010): Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*. 214:655-667.
47. Sridharan D, Levitin DJ, Menon V (2008): A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*. 105:12569-12574.
48. Lindquist KA, Satpute AB, Wager TD, Weber J, Barrett LF (2015): The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*. 26:1910-1922.
49. Phan KL, Wager T, Taylor SF, Liberzon I (2002): Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*. 16:331-348.
50. Hamilton JP, Etkin A, Furman DJ, Lemus MG, Johnson RF, Gotlib IH (2012): Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural response data. *American Journal of Psychiatry*. 169:693-703.
51. Ma Y (2015): Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. *Molecular psychiatry*. 20:311.
52. Harmer CJ, O'Sullivan U, Favaron E, Massey-Chase R, Ayres R, Reinecke A, et al. (2009): Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am J Psychiatry*. 166:1178-1184.
53. Warren MB, Pringle A, Harmer CJ (2015): A neurocognitive model for understanding treatment action in depression. *Philos Trans R Soc Lond B Biol Sci*. 370:20140213.
54. Harmer CJ, Goodwin GM, Cowen PJ (2009): Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *The British Journal of Psychiatry*. 195:102-108.
55. Beck AT (2008): The evolution of the cognitive model of depression and its neurobiological correlates.
56. McGrath CL, Kelley ME, Dunlop BW, Holtzheimer Iii PE, Craighead WE, Mayberg HS (2014): Pretreatment Brain States Identify Likely Nonresponse to Standard Treatments for Depression. *Biological Psychiatry*. 76:527-535.
57. Konarski JZ, Kennedy SH, Segal ZV, Lau MA, Bieling PJ, McIntyre RS, et al. (2009): Predictors of nonresponse to cognitive behavioural therapy or venlafaxine using glucose metabolism in major depressive disorder. *Journal of psychiatry & neuroscience: JPN*. 34:175.
58. Siegle GJ, Carter CS, Thase ME (2006): Use of FMRI to predict recovery from unipolar depression with cognitive behavior therapy. *American Journal of Psychiatry*. 163:735-738.
59. Siegle GJ, Thompson WK, Collier A, Berman SR, Feldmiller J, Thase ME, et al. (2012): Toward clinically useful neuroimaging in depression treatment: prognostic utility of subgenual cingulate activity for determining depression outcome in cognitive therapy across studies, scanners, and patient characteristics. *Archives of General Psychiatry*. 69:913-924.
60. Dunlop BW, Rajendra JK, Craighead WE, Kelley ME, McGrath CL, Choi KS, et al. (2017): Functional Connectivity of the Subcallosal Cingulate Cortex

And Differential Outcomes to Treatment With Cognitive-Behavioral Therapy or Antidepressant Medication for Major Depressive Disorder. *Am J Psychiatry*. 174:533-545.
61. Roiser JP, Elliott R, Sahakian BJ (2012): Cognitive Mechanisms of Treatment in Depression. *Neuropsychopharmacology*. 37:117-136.
62. Gyurak A, Gross JJ, Etkin A (2011): Explicit and implicit emotion regulation: a dual-process framework. *Cognition and emotion*. 25:400-412.
63. Etkin A, Egner T, Kalisch R (2011): Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in cognitive sciences*. 15:85-93.
64. Altman DG (1985): Comparability of randomised groups. *The Statistician*.125-136.
65. Schulz KF, Altman DG, Moher D (2010): CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC medicine*. 8:18.
66. Kazdin AE (2007): Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol*. 3:1-27.
67. Holmes EA, Ghaderi A, Harmer CJ, Ramchandani PG, Cuijpers P, Morrison AP, et al. (2018): The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry*. 5:237-286.