Research paper

Examining raphe-amygdala structural connectivity as a biological predictor of SSRI response

Rajapillai L.I. Pillai a, Huang Chuan a,b,*, Andrew LaBella a, Zhang Mengru d, Yang Jie e, Madhukar Trivedi f, Myrna Weissman g, Patrick McGrath g, Maurizio Fava h, Benji Kurian f, Crystal Cooper f, Melvin McInnis i, Maria A. Oquendo j, Diego A. Pizzagalli h, Ramin V. Parsey a, Christine DeLorenzo a,k

a Department of Psychiatry, Stony Brook University, Stony Brook, NY, United States
b Department of Radiology, Stony Brook University, Stony Brook, NY, United States
c Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY, United States
d Department of Family, Population, & Preventive Medicine, Stony Brook University, Stony Brook, NY, United States
e Department of Psychiatry, University of Texas Southwestern Medical Center, United States
f Department of Psychiatry, College of Physicians and Surgeons, Columbia University and the New York Psychiatric Institute, United States
g Department of Psychiatry, Harvard Medical School, United States
h Department of Psychiatry, University of Michigan, United States
i Department of Psychiatry, University of Pennsylvania, United States
j Department of Psychiatry, Molecular Imaging and Neuropathology Division, Columbia University, New York, NY, United States

ARTICLE INFO

Keywords:
Diffusion tensor imaging
Amygdala
Raphe nucleus
Fractional anisotropy
Remission
SSRI

ABSTRACT

Background: Our lab has previously found that structural integrity in tracts from the raphe nucleus (RN) to the amygdala, measured by fractional anisotropy (FA), predicts remission to selective serotonin reuptake inhibitors (SSRIs) in major depressive disorder (MDD). This could potentially serve as a biomarker for remission that can guide clinical decision-making. To enhance repeatability and reproducibility, we replicated our study in a larger, more representative multi-site sample.

Methods: 64 direction DTI was collected in 144 medication-free patients with MDD from the Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care (EMBARC) study. We performed probabilistic tractography between the RN and bilateral amygdala and hippocampus and calculated weighted FA in these tracts. Patients were treated with either sertraline or placebo, and their change in Hamilton Depression Rating Scale (HDRS) score reported. Pretreatment weighted FA was compared between remitters and non-remitters, and correlation between FA and percent change in HDRS score was assessed. Exploratory moderator and voxel analyses were also performed.

Results: Contrary to our hypotheses, FA was greater in nonremitters than in remitters in RN-left and right amygdala tracts ($p = 0.02$ and $0.01$, respectively). Pretreatment FA between the raphe and left amygdala correlated with greater, not reduced, HDRS ($r = 0.18$, $p = 0.04$). This finding was found to be greater in the placebo group. Moderator and voxel analyses yielded no significant findings.

Conclusions: We found greater FA in nonremitters between the RN and amygdala than in remitters, and a correlation between FA and symptom worsening, particularly with placebo. These findings may help reveal more about the nature of MDD, as well as guide research methods involving placebo response.

https://doi.org/10.1016/j.jad.2019.05.055
Received 29 September 2018; Received in revised form 18 April 2019; Accepted 27 May 2019
Available online 28 May 2019

© 2019 Elsevier B.V. All rights reserved.
1. Introduction

Major depressive disorder (MDD) is a heterogeneous mental disorder with complex etiology and a variety of negative outcomes. Despite being one of the most prevalent diseases globally, extensive research and clinical treatment trials have fallen short of consistently helping patients achieve remission (Katzman et al., 2014). Selective Serotonin Reuptake Inhibitors (SSRIs), the most common MDD treatment, only produce remission in one out of three patients (Rush et al., 2006). Ineffective SSRI treatment can often cause additional impairments due to common side effects such as insomnia, sexual dysfunction, and anxiety (Lam et al., 2012). Moreover, SSRIs often take four weeks or longer to exert antidepressant effects (Montgomery, 1997). Therefore, the ability to accurately predict response prior to treatment would help determine whether an SSRI is appropriate, or whether another form of therapy should be pursued. Such accuracy may result in higher treatment adherence, less patients lost to follow-up, and higher success rates in treating MDD.

A non-invasive, in vivo option that can be used for this purpose is Diffusion Tensor Imaging (DTI), a form of Magnetic Resonance Imaging (MRI) that measures the diffusion of water molecules in the brain (Troedel and Amico, 2014). One variable extracted from DTI is Fractional Anisotropy (FA), which is derived from the calculated mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) of water molecules (Alexander et al., 2007; Soares et al., 2013). FA is a measure of the directional dependence of water diffusivity, with 0 corresponding to isotropic diffusion in all directions and 1 indicating that water diffuses completely across one axis (i.e., completely anisotropic) (O’Donnell and Westin, 2011). It is thought that high FA corresponds to dense tracts of myelinated WM axons (Du et al., 2014). Low FA values may reflect decreased myelin sheath integrity (Heckel et al., 2015), although it can be influenced by factors such as noise, artifacts, and crossing fibers. DTI can also be used to estimate the location of fiber tracts from one region (seed) to another (target).

Pretreatment FA differences between those who recover with treatment (remitters) and those who do not (non-remitters) may reflect baseline differences in structural integrity in brain regions implicated in treatment efficacy. In line with this assumption, studies have found associations between treatment response (including SSRIs, ketamine, and a selective norepinephrine reuptake inhibitor (SNRI)) and pretreatment FA in regions such as the cingulum, stria terminalis (Grieve et al., 2016; Korgaonkar et al., 2014), and frontal cortex (Vasavada et al., 2016). These regions are connected to areas known to be affected in MDD, such as the subgenual anterior cingulate cortex (Vergani et al., 2016) and amygdala (Kruger et al., 2015). However, a clinically viable marker for treatment response using DTI has not yet been validated and has only been observed in small patient samples with small effect sizes.

SSRIs’ initial chief mechanism of action involves blocking the serotonin transporter (5-HTT), preventing reuptake of serotonin and prolonging its action in the synapse (Fuller and Wong, 1977). Therefore, the efficacy of SSRIs can be affected by the overall health of the serotonin system. The principal source of serotonin in the central nervous system is the raphe nucleus (RN) (Hornung, 2003), found in the midbrain,pons, and medulla. Serotonergic fibers project from this region to nearly all areas of the brain. Therefore, FA between the RN and these regions can serve as an indication of serotonergic integrity. The amygdala is a logical choice of target: it is known to play a central role in several areas. Pretreatment FA differences due to common side effects such as insomnia, sexual dysfunction, and anxiety (Lam et al., 2012). Moreover, SSRIs often take four weeks or longer to exert antidepressant effects (Montgomery, 1997). Therefore, the ability to accurately predict response prior to treatment would help determine whether an SSRI is appropriate, or whether another form of therapy should be pursued. Such accuracy may result in higher treatment adherence, less patients lost to follow-up, and higher success rates in treating MDD.

2. Methods and materials

2.1. Participants

Data for all participants were taken from the EMBARC project, which enrolled subjects at University of Michigan, Columbia University Medical Center, Massachusetts General Hospital, and University of Texas Southwestern Medical Center. EMBARC was designed to find pretreatment biomarkers of treatment response. All four sites had their protocols reviewed and approved by their Institutional Review Boards, and all participants provided written informed consent.

Eligibility for each participant was determined based on a psychiatric assessment, review of medical history, chart review, clinical interview, physical examination, routine blood tests, pregnancy test, and urine toxicology. Detailed inclusion/exclusion criteria can be found in Trivedi et al. (2016). Briefly, inclusion criteria comprised: (1) age 18–65; (2) met criteria for a major depressive episode according to the Structured Clinical Interview for the DSM IV (First et al., 1995); (3) scored at least 14 on the Quick Inventory of Depressive Symptoms (QIDS-SR) (Rush et al., 2003); and (4) had early onset (< 30 years old) and chronic (> 2 years in duration) MDD or recurrent MDD (2 or more recurrences). The study comprised 309 participants. Of these, 270 had structural MRI images that passed stringent quality control measures and 172 of these had DTI images that passed strict quality control measures, examined by trained technicians (see below). To ensure that remission was clinically significant, the sample was further restricted to patients with a Hamilton Depression Rating Scale (HDRS) 17-item (Hamilton, 1960) score of 15 or greater at baseline, as there is evidence that response correlates with depression severity and in mild depression there is a failure to differentiate placebo from drug response (Fournier et al., 2010). This final set comprised 144 patients (Table 1).
Participants were classified as remitters if their final HDRS scores were less than or equal to 7 after 8 weeks of treatment on either sertraline or placebo; if these criteria were not met, they were classified as non-remitters. In addition, patients were considered responders if their HDRS did not.

2.2. MRI acquisition

MRI scans were performed at Columbia University Medical Center (GE Signa HDx), Massachusetts General Hospital (Siemens TrioTim), University of Michigan (Phillips Ingenia), or University of Texas Southwestern Medical Center (Phillips Achieva). Structural MRIs were acquired using the following parameters: TR (repetition time): 5.9–8.2 ms, TE (echo time): 2.4–3.7 ms, Flip Angle: 9°–12°, voxel dimensions: 1 mm × 1 mm × 1 mm, acquisition matrix: 240 × 240, 256 × 256 or 256 × 256, acceleration factor: 2, and 174–178 sagittal slices. Scan parameters can be found in Supplementary Table 1, and have been published previously (Fournier et al., 2017). Diffusion images were acquired using a single-shot EPI (echo planar imaging) sequence. Scan parameters were as follows: TR = 8310–9500 ms, TE = 95–96.3 ms, Flip Angle = 90°, voxel dimensions = 2.5 mm × 2.5 mm × 2.5 mm or 1.9 mm × 1.9 mm × 2.5 mm, acquisition matrix = 96 × 96, b value = 1000 s/mm², and 64 diffusion directions with 1 or 5 non-weighted images (b = 0). DTI acquisition time was approximately 10 min. These parameters were selected prior to data acquisition to standardize acquisition between scanners as much as possible (Zhu et al., 2011). To adjust for any residual effects, site was also included as a covariate.

2.3. MRI processing

T1-weighted structural MRI images were assessed for quality and processed at Stony Brook University by trained technicians. Quality inspection included checking for common artifacts such as field inhomogeneity and poor skull stripping. Images were approved if the amygdala, hippocampus, and brainstem (all regions used in this work) were devoid of artifacts that would interfere with region identification.

Structural MRIs were processed using Freesurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/). Amygdala and hippocampus ROIs were derived from Freesurfer’s atlases. The RN ROI was derived from a previous sample from our laboratory, in which 52 healthy controls were scanned using the Positron Emission Tomography (PET) tracer [11C]-WAY100635, a serotonin 1A antagonist (Forster et al., 1995). Voxel binding maps were calculated for each subject and warped into MNI space. A threshold binding potential of 18 mL/cc was then applied to the midbrain region of these images to differentiate the RN, as this area has higher serotonin 1A binding than surrounding regions. This RN ROI was then inverse-warped to individual MRIs using Advanced Normalization Tools (ANTs) (Avants et al., 2011). We have used this method to delineate the RN previously (Delorenzo et al., 2013).

2.4. DTI processing

DTI images were also assessed for quality and processed at Stony Brook University. Diffusion weighted images were assessed for quality by trained technicians to check for artifacts such as venetian blind, gradient-wise motion, ghost, ringing, and slice-wise intensity artifacts (Liu et al., 2010). Images were approved if the ventral area encompassing the amygdala, hippocampus, and brainstem was devoid of

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Total</th>
<th>Placebo</th>
<th>Sertraline</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>All participants (74 placebo vs 70 sertraline)</td>
<td>37.27 ± 13.71</td>
<td>36.38 ± 13.36</td>
<td>38.21 ± 14.10</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>90 (62.50%)</td>
<td>43 (47.78%)</td>
<td>47 (52.22%)</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>54 (37.50%)</td>
<td>31 (57.41%)</td>
<td>23 (42.59%)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>American Indian/Alaska Native</td>
<td>1 (0.69%)</td>
<td>1 (100.00%)</td>
<td>0 (0.00%)</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>11 (7.64%)</td>
<td>7 (63.64%)</td>
<td>4 (36.36%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black or African American</td>
<td>24 (16.67%)</td>
<td>12 (50.00%)</td>
<td>12 (50.00%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>9 (6.25%)</td>
<td>5 (55.56%)</td>
<td>4 (44.44%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>99 (68.75%)</td>
<td>49 (49.49%)</td>
<td>50 (50.51%)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>No</td>
<td>120 (83.33%)</td>
<td>61 (50.83%)</td>
<td>59 (49.17%)</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>24 (16.67%)</td>
<td>13 (54.17%)</td>
<td>11 (45.83%)</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>CU</td>
<td>46 (31.94%)</td>
<td>22 (47.83%)</td>
<td>24 (52.17%)</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>31 (21.53%)</td>
<td>18 (58.06%)</td>
<td>13 (41.94%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TX</td>
<td>43 (29.86%)</td>
<td>23 (53.49%)</td>
<td>20 (46.51%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>24 (16.67%)</td>
<td>11 (45.83%)</td>
<td>13 (54.17%)</td>
<td></td>
</tr>
<tr>
<td>MDD severityd</td>
<td>High</td>
<td>74 (51.39%)</td>
<td>36 (48.65%)</td>
<td>38 (51.35%)</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>70 (48.61%)</td>
<td>38 (54.29%)</td>
<td>32 (45.71%)</td>
<td></td>
</tr>
<tr>
<td>MDD chronicityd</td>
<td>Chronic</td>
<td>73 (50.69%)</td>
<td>34 (46.58%)</td>
<td>39 (53.42%)</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Non chronic</td>
<td>71 (49.31%)</td>
<td>40 (53.64%)</td>
<td>31 (46.66%)</td>
<td></td>
</tr>
<tr>
<td>Remitter</td>
<td>No</td>
<td>91 (63.19%)</td>
<td>49 (53.85%)</td>
<td>42 (46.15%)</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>53 (36.81%)</td>
<td>25 (47.17%)</td>
<td>28 (52.83%)</td>
<td></td>
</tr>
<tr>
<td>Week 0 HDRS score</td>
<td>All participants</td>
<td>19.83 ± 3.59</td>
<td>19.55 ± 3.49</td>
<td>20.11 ± 3.70</td>
<td>0.35</td>
</tr>
<tr>
<td>Week 8 HDRS score</td>
<td>All participants</td>
<td>−8.38 ± 7.36</td>
<td>−7.62 ± 7.30</td>
<td>−9.17 ± 7.40</td>
<td>0.21</td>
</tr>
<tr>
<td>% Change in HDRS score</td>
<td>All participants</td>
<td>−42 ± 37</td>
<td>−40 ± 38</td>
<td>−45 ± 36</td>
<td>0.43</td>
</tr>
<tr>
<td>Weighted FA between raphe and left amygdala</td>
<td>All participants</td>
<td>0.42 ± 0.04</td>
<td>0.42 ± 0.03</td>
<td>0.43 ± 0.04</td>
<td>0.43</td>
</tr>
<tr>
<td>Weighted FA between raphe and right amygdala</td>
<td>All participants</td>
<td>0.43 ± 0.04</td>
<td>0.42 ± 0.04</td>
<td>0.43 ± 0.04</td>
<td>0.21</td>
</tr>
<tr>
<td>Weighted FA between raphe and left hippocampus</td>
<td>All participants</td>
<td>0.44 ± 0.05</td>
<td>0.43 ± 0.05</td>
<td>0.44 ± 0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>Weighted FA between raphe and right hippocampus</td>
<td>All participants</td>
<td>0.45 ± 0.05</td>
<td>0.44 ± 0.05</td>
<td>0.46 ± 0.05</td>
<td>0.03*</td>
</tr>
<tr>
<td>Concurrent medication use</td>
<td>Yes</td>
<td>78 (54.17%)</td>
<td>39 (50%)</td>
<td>39 (50%)</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>66 (45.83%)</td>
<td>35 (50.53%)</td>
<td>31 (49.67%)</td>
<td></td>
</tr>
</tbody>
</table>

a Total displayed: mean ± SD for continuous variables; N (%) for categorical variables.
b For continuous variables, p-values were calculated from Welch’s t-tests; for categorical variables, exact p-values were calculated from the Monte Carlo simulation. All p-values are calculated between placebo and sertraline groups.
c Defined by HDRS ≥ 20 and QIDS-SR of ≥ 16.
d Chronically defined as ≥ 2 years.
e Not significant after multiple comparisons correction.
artifacts that would interfere with probabilistic tractography. A relatively large number of studies were excluded in this process (58) for reasons stated above, due in part to the fact that DTI is sensitive to small motions. When demographics, HDRS scores, and remission status was compared between patients with approved DTI and HDRS score > 15 (144 total) and those with HDRS score > 15 who did not have approved DTI, no significant differences were found.

Images were corrected for motion and gradient coil-induced distortions using the eddy current correction routine in FSL (FMRI Software Library, http://www.fmrib.ox.ac.uk/fsl/). Nonbrain tissue was removed through FSL’s Brain Extraction Tool. After this, FA was computed in Camino (http://web4.cs.ucl.ac.uk/research/medic/camino/) using the least squares fit diffusion tensor with nonlinear optimization using a Levenburg–Marquardt algorithm, constrained to positive values by fitting its Cholesky decomposition (Alexander and Barker, 2005).

Probabilistic Tractography was performed using the FMRI Diffusion Toolbox (FDT) (Behrens et al., 2007). FDT computes streamlines through each voxel by repeated sampling from each principal diffusion direction. This method yields the probability of connections from the seed (RN) to post-synaptic targets in the brain (amygdala and hippocampus). FDT was run with 5000 samples, a maximum of 2000 steps per sample, a step length of 0.5 mm, and a tract curvature threshold of 0.2 mm. We used this to compute the weighted average FA within the subject-specific tractography-defined tracts from the RN to the amygdala and hippocampus (Fig. 1). To include only voxels with a high probability of being in the defined tract, weighted average FA was calculated by multiplying voxel-based FA measures by the probability of connection at each voxel, summing the products, and dividing by the sum of the probabilities (Bonnelle et al., 2012; Hagler et al., 2009; Hua et al., 2008). Similar weighting procedures were used for mean, radial, and axial diffusivity.

2.5. Statistics

All p-values for demographics comparisons were derived using Monte-Carlo simulations (Agresti et al., 1979). A linear mixed model was used to compare weighted FA between remitters and non-remitters after covarying for age, sex, and site, and treatment (placebo vs. sertraline). The two-way interaction term among the remission measures and regions was considered in order to account for possible different region-specific relationships in the left and right amygdala and hippocampus. This model yields many more samples to estimate residual errors, increasing detection power and allowing us to measure both region-specific effects and group differences between regions. Akaike Information Criteria (AIC) was used when selecting dependence structures for modeling the correlation among imaging measures from different regions but the same patient. Possible covariance structures considered were unstructured (UN), compound symmetry (CS), heterogeneous CS (CSH), Toeplitz (TOEP); CSH was selected according to AIC. Similar models were used to test mean, radial, and axial diffusivity. To test for correlations of FA with decrease in MDD symptoms, a similar linear mixed model was used, but with percent change in HDRS score as an outcome measure, defined as:

$$x = \frac{\text{Week 8 Score} - \text{Week 0 Score}}{\text{Week 0 Score}}$$

By this formulation, an $x$ value of 0 indicates no change, a value $< 0$ indicates improvement, and a value $> 0$ indicates worsening. Note that this is the opposite as our pilot study (Deforenzo et al., 2013), where positive values indicate improvement—we performed the analysis this way to denote a change from baseline in any direction. Partial correlation coefficients were calculated to quantify the correlation between FAs and decrease in MDD symptoms after removing possible confounding effects of age, sex and site through the linear mixed model.

Multiple linear regression models and multiple logistic regression models were also fitted to test if tracts from the raphe to the left amygdala or right amygdala are treatment response moderators by examining the interaction term between FA values and treatment groups. Possible covariates included age, sex and site, which were significantly related to treatment response in the univariate data analysis. Statistical significance was set at 0.05. Mixed model analysis was performed using SAS 9.4 (SAS Institute Inc., Cary, NC). Multiple comparisons were corrected for via the Bonferroni method for the left and right amygdala.

2.6. Voxel analysis

An exploratory voxel analysis of FA differences was carried out using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) — this test compared FA voxel values across all subjects in standard space using a full factorial model with age, sex, and site added as covariates. Cluster size was thresholded at four voxels, and family-wise error correction was used. For this analysis, only participants for whom MRI and DTI processing for all regions had been approved were included—this yielded a smaller sample than our a priori analyses, which only required approval of amygdala, hippocampus, and RN. The sample comprised 98 participants. A significance threshold of $p < 0.05$ family-wise error corrected threshold was used.

![Fig. 1. Coronal (A), Sagittal (B), and axial (C) images of a representative tract (red) from the raphe nucleus (green) to the amygdala (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
3. Results

3.1. Fractional anisotropy in remitters and non-remitters

All comparisons were performed across both placebo and sertraline groups unless stated otherwise. Groupwise comparisons of remitters vs. non-remitters revealed no significant differences in fractional anisotropy (FA) in tracts from the raphe nucleus (RN) to both left and right amygdala ($p = 0.02$ and $0.01$, respectively, Bonferroni corrected; see Fig. 2). No significant difference was found in tracts from RN to left ($p = 0.16$, uncorrected) or right ($p = 0.82$, uncorrected) hippocampus. Average difference in FA between groups was 0.02 in the left and right amygdala, 0.01 in the left hippocampus, and less than 0.01 in the right hippocampus (Table 2)—in all cases, remitters had lower FA than non-remitters.

Similar analysis performed between responders and non-responders revealed no significant differences in tracts from the RN to the left or right amygdala ($p = 0.29$ and 0.18, respectively, uncorrected), or to the left or right hippocampus ($p = 0.69$ and $>0.99$, respectively, uncorrected, Supplementary Table 2).

3.2. Mean, radial, and axial diffusivity between remitters and non-remitters

Neither mean diffusivity, nor radial diffusivity, nor axial diffusivity differed significantly between remitters and non-remitters, or between responders and non-responders in any of the tracts examined. Moreover, all differences in means for these measures were less than 0.00001 (Supplemental Tables 3–5).

3.3. Moderator analysis of left amygdala and right amygdala

No interaction term between FA values and treatment group (sertraline vs. placebo) was significant, i.e., there is no evidence that FA influences the relationship between treatment group and remission

3.4. Correlation of fractional anisotropy with symptom improvement

When the correlation between average FA in the defined tracts and percent change in HDRS was examined, a significant positive correlation was found in average FA of tracts connecting the RN to the left amygdala ($r = 0.39$, see Fig. 3). However, these independent correlations were not significantly different ($Z = 1.59$, ns).

3.5. Exploratory voxel analysis

An exploratory voxel analysis found no significant differences in FA between remitters and non-remitters, including in previously delineated regions of frontal cortex, cingulum, and stria terminalis (Alexopoulos et al., 2002, 2008; Grieve et al., 2016; Korgaonkar et al., 2014; Taylor et al., 2008).

4. Discussion

Successful replication in neurophysiological research has often proven elusive, and there is a broad effort in the scientific community to address this (Kappenman and Keil, 2017). As part of this effort, we sought to replicate our pilot study in a larger, more representative sample. Similar to our previous study, we found a significant difference.

<table>
<thead>
<tr>
<th>Seed to target pair</th>
<th>Partial correlation coefficient</th>
<th>95% CI lower</th>
<th>95% CI upper</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raphe to left amygdala</td>
<td>0.18</td>
<td>0.01</td>
<td>0.33</td>
<td>0.04*</td>
</tr>
<tr>
<td>Raphe to right amygdala</td>
<td>0.09</td>
<td>−0.08</td>
<td>0.25</td>
<td>0.42</td>
</tr>
<tr>
<td>Raphe to left hippocampus</td>
<td>0.05</td>
<td>−0.12</td>
<td>0.21</td>
<td>0.44</td>
</tr>
<tr>
<td>Raphe to right hippocampus</td>
<td>−0.05</td>
<td>−0.22</td>
<td>0.11</td>
<td>0.69</td>
</tr>
</tbody>
</table>

HDRS: Hamilton Depression Rating Scale, CI: confidence interval. $P$-values are from linear mixed model and are Bonferroni corrected for left and right amygdala. Partial correlation coefficients were calculated from linear mixed model after adjusting for age, sex and study site. *$p < 0.05$. 
(p = 0.59 and 0.95 for tracts terminating in the left and right amygdala, respectively). Similarly, there was no significant treatment interaction in the relative change in HDRS score for left or right amygdala ($p = 0.27$ and 0.91, respectively).

Table 3

Correlation of fractional anisotropy with percent change in HDRS.

Table 2

Comparison of DTI measures between remitters and non-remitters.  

<table>
<thead>
<tr>
<th>Seed to target pair</th>
<th>Average FA (Remitters)</th>
<th>Average FA (Non-remitters)</th>
<th>Average FA Difference</th>
<th>95% CI lower</th>
<th>95% CI upper</th>
<th>Cohen’s $D$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raphe to left amygdala</td>
<td>0.41</td>
<td>0.43</td>
<td>−0.02</td>
<td>−0.03</td>
<td>&gt; −0.01</td>
<td>0.39</td>
<td>0.02*</td>
</tr>
<tr>
<td>Raphe to right amygdala</td>
<td>0.41</td>
<td>0.43</td>
<td>−0.02</td>
<td>−0.03</td>
<td>&lt; −0.01</td>
<td>0.54</td>
<td>0.01*</td>
</tr>
<tr>
<td>Raphe to left hippocampus</td>
<td>0.43</td>
<td>0.44</td>
<td>−0.01</td>
<td>−0.03</td>
<td>&lt; 0.01</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td>Raphe to right hippocampus</td>
<td>0.45</td>
<td>0.45</td>
<td>&lt; −0.01</td>
<td>−0.02</td>
<td>0.01</td>
<td>0.10</td>
<td>0.82</td>
</tr>
</tbody>
</table>

* DTI: diffusion tensor imaging, FA: fractional anisotropy, CI: confidence interval. $P$-values are from linear mixed model and are Bonferroni corrected for left and right amygdala. Differences and $p$-values were calculated from linear mixed model after adjusting for age, gender and study site.

### Notes
- $^a$ Difference is calculated as mean FA(remitters) – mean FA(non-remitters).
- $^b$ $p < 0.05$. 

Fig. 2. Boxplot of residual fractional anisotropy (FA) after taking age, sex, and site into account. Group means are demarcated by the “X”. Right amygdala is plotted as this is the region with the greatest difference between remitters and non-remitters.
in FA between remitters and non-remitters, but we found the opposite pattern; rather than positively correlating with symptom improvement, fractional anisotropy (FA) in the tracts between the raphe nucleus (RN) and amygdala positively correlated with symptom worsening, and non-remitters had higher FA between the raphe and amygdala than remitters. A reversal such as this may have been due to the effect

Fig. 3. Output of linear mixed models relating change in symptom severity to fractional anisotropy (FA) between the raphe nucleus and left amygdala for all participants (A), placebo group only (B), and sertraline group only (C). The placebo group drives the overall correlation. None of other correlations examined were significantly different. HDRS-17: Hamilton Depression Rating Scale, 17 Item.
size—with a correlation coefficient of 0.18, our previous study of 18 participants was likely underpowered and may have found a positive correlation by chance. As our present study has nearly ten times the number of participants, our statistical power is greatly improved. There are some differences in study design from our preliminary analysis: unlike our previous study, this was a placebo-controlled trial (although examination of the SSRI arm alone revealed no significant association with remission or symptom improvement). This also represents a difference from most clinical scenarios, where a patient is aware of what medicine they are receiving. In addition, while our previous study used a combined midbrain and RN ROI, we used the raphe by itself for our current analysis as our improved DTI resolution allowed us to explore this region. When we repeated the analysis with the combined ROI, however, we replicated our current results (data not shown). It is important to note that DTI cannot provide contrast for brainstem nuclei in of itself—therefore the integrity of the region depends on its quality of delineation from PET studies. We have used this method previously (Delorenzo et al., 2013). While an MRI-based method would be valuable, this has not yet been developed. One advantage of our PET based method is that it allows the raphe delineation to be subject-specific.

While this finding is in contrast to studies showing higher pre-treatment FA in frontal cortex in remitters (Alexopoulos et al., 2002, 2008) our results resemble those of a similar analysis performed in frontal cortex in MDD patients—in that study, higher FA in the superior frontal gyrus and anterior cingulate was associated with a poorer outcome to sertraline (Taylor et al., 2008). Another study found that higher FA in the stria terminalis predicted non-remission, while higher FA in the cingulum predicted remission (Korgaonkar et al., 2014). This was replicated when the authors used the ratio of cingulum to stria terminalis binding to predict remission vs. non-remission (Grieve et al., 2016). While we were not able to replicate these findings in our own voxel analysis, the stria terminalis is particularly relevant in that it contains efferent fibers from the amygdala (Wakana et al., 2004). Given findings of increased amygdala activity in MDD (Sheline et al., 2001), it may be that connectivity to this region is detrimental to treatment. However, no significant differences in FA were observed between remitter and non-remitter in our exploratory whole-brain voxel analysis, suggesting that the effect size may not be large enough to withstand multiple corrections in an exploratory analysis. Indeed, the difference between groups in FA between the RN and amygdala was only 0.02.

While these findings cannot be applied on an individual basis at this time, our study reveals group-level findings that may give some insight into pathophysiology. Our previous hypothesis was that patients with higher FA would show greater clinical improvement, possibly by capitalizing on more robust serotonergic pathways. Our current results suggest that the opposite may be true. One possibility for this may be that SSRIs are more beneficial for patients who do not have robust serotonergic fibers, and that these drugs act to enhance these pathways. Longitudinal analysis of FA in these pathways would be needed to support this hypothesis.

A previous meta-analysis examined factors that predicted remission with placebo—these included patients with less severe MDD (lower Hamilton Depression Rating Scale scores), younger age, less anxiety, and shorter MDD episode duration (Nelson et al., 2012). Our placebo group happened to be slightly younger, have slightly lower HDRS scores, and have a smaller proportion with chronic MDD (Table 1). Despite the fact that none of these reached significance, they may have individually contributed to the lack of differences in outcomes among treatment groups.

Despite the moderate effect sizes, the absolute difference between groups was small and the predictive value of FA in these regions was only slightly better than chance (area under the curve using ten rounds of five-fold cross-validation was 0.57, while random chance would be 0.5—data not shown). In addition, there appears to be no moderating effect of FA either on remission or on changes in depression severity. Therefore, these findings can be taken as negative on a clinical level, and do not replicate previous results. It should be noted that these findings were negative with and without covariates (data not shown). This parallels previous studies reporting negative findings in FA between patients with MDD and healthy controls (Choi et al., 2014; Olvet et al., 2015). As this study had a large sample size, it raises important limitations about the ability of FA to influence clinical decisions.

Given that well-powered DTI studies have failed to find clinically significant differences between healthy controls and patients with MDD (Choi et al., 2014) or, in our sample, between remitters and non-remitters, research into biomarkers of diagnosis and prognosis may require further design considerations. For example, MDD is a heterogeneous disorder. Such heterogeneity may have confounded the search for a single marker across all subjects. Therefore, a more symptomatic approach, such as studying clinical subtypes or research domain criteria (RDoC, a classification system based on translational neurobiology and observable behavior as opposed to self-reported categories of symptoms) (Cuthbert and Insel, 2013) may reduce these confounding effects. Moreover, many patients were taking concurrent medications. While our study prohibited psychoactive medication in our study, it is possible that other classes of medications had effects on DTI measures that have not yet been studied. In addition, as mentioned, DTI cannot selectively identify neurons by their emitted neurotransmitter or by the direction of transmission. This limits our interpretation of FA findings. While there is currently no in vivo method to select for tracts based on their primary neurotransmitter, a combined PET/DTI study examining binding potential in a seed and target region in conjunction with tract integrity between them may yield valuable information. However, it is important to note that across the EMBARC study a statistically significant difference in outcome between the placebo and sertraline groups in clinical outcome was not observed (Webb et al., 2019), which may have contributed to the lack of predictive power of our study.

In conclusion, in this large multi-site study, we found greater FA in non-remitters between the RN and amygdala than in remitters. Moreover, we found a correlation between FA in the tracts connecting the RN to the left amygdala and symptom worsening. While this measure may not be useful on an individual basis with current technology, these results may help reveal more about the nature of MDD. The use of symptom-specific, RDoC, and multimodal approaches may bring us closer to finding a clinically-useful prognostic marker.

Financial disclosures


In the past two years, Dr. Myrna Weissman received funding from the National Institute of Mental Health (NIMH), the Sackler Foundation, the Templeton Foundation; and receives royalties from the Oxford University Press, Perseus Press, the American Psychiatric
Association Press, and MultiHealth Systems.

Dr. Patrick McGrath has received funding from the National Institute of Mental Health, New York State Department of Mental Hygiene, Research Foundation for Mental Hygiene (New York State), Forest Research Laboratories, Sunovion Pharmaceuticals, and Naurex Pharmaceuticals (now Allergan).

Dr. Maurizio Fava has received research support from Abbott Laboratories; Alkermes, Inc.; American Cyanamid; Aspect Medical Systems; AstraZeneca; Avanir Pharmaceuticals; BioResearch; BrainCells Inc.; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon; Clintara, LLC; Cerecor; Covance; Covidien; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Forest Pharmaceuticals, Inc.; Ganeden Biotech, Inc.; GlaxoSmithKline; Harvard Clinical Research Institute; Hoffman-LaRoche; Icon Clinical Research; i3 Innovus/Ingenix; Janssen R&D, LLC; Jed Foundation; Johnson & Johnson Pharmaceutical Research & Development; Lichtwer Pharma GmbH; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante; Metylation Sciences Inc.; National Alliance for Research on Schizophrenia & Depression (NARSAD); National Center for Complementary and Alternative Medicine (NCCAM); National Institute of Drug Abuse (NIDA); National Institute of Mental Health (NIMH); Neuralstem, Inc.; Novartis AG; Organon Pharmaceuticals; PamLab, LLC.; Pfizer Inc.; Pharmacia-Upjohn; Pharmaceutical Research Associates., Inc.; Pharmavite® LLC; PharmoRx Therapeutics; Photothera; Reckitt Benckiser; Roche Pharmaceuticals; RCT Logic, LLC (formerly Clinical Trials Solutions, LLC); Sanofi-Aventis US LLC; Shire, Solvay Pharmaceuticals, Inc.; Stanley Medical Research Institute (SMRI); Synthelabo; Tal Medical; Wyeth-Ayerst Laboratories; he has served as advisor or consultant to Abbott Laboratories; Acadia; Afectis Pharmaceuticals AG; Alkermes, Inc.; Amaryn Pharma Inc.; Aspect Medical Systems; AstraZeneca; Auspex Pharmaceuticals; Avanir Pharmaceuticals; AXSOME Therapeutics; Bayer AG; Best Practice Project Management, Inc.; Biogen; BioMarin Pharmaceuticals, Inc.; Biowall Corporation; BrainCells Inc; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon, Inc.; Cerecor; CNS Response, Inc.; Compellis Pharmaceuticals; Cypress Pharmaceutical, Inc.; DiagnoSearch Life Sciences (P) Ltd.; Dinippon Sumitomo Pharma Co. Inc.; Dov Pharmaceuticals, Inc.; Edgemont Pharmaceutical, Inc.; Eisai Inc.; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; ePharmaSolutions; EPIX Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Fabre-Kramer Pharmaceuticals, Inc.; Forest Pharmaceuticals, Inc.; Forum Pharmaceuticals, Inc.; GenOmind, LLC; GlaxoSmithKline; Grunenthal GmbH; i3 Innovus/Ingenix; Intracellular; Janssen Pharmaceuticals; Jazz Pharmaceuticals, Inc.; Johnson & Johnson Pharmaceutical Research & Development, LLC; Knoll Pharmaceutical Corporations; Labopharm Inc.; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante, Inc.; Merck & Co., Inc.; MSI Metylation Sciences, Inc.; Naurex, Inc.; Nestle Health Sciences; Neuralstem, Inc.; Neuroetics, Inc.; NextWave Pharmaceuticals; Novartis AG; Nutrition 21; Orexigen Therapeutics, Inc.; Organon Pharmaceuticals; Osmotica; Otsuka Pharmaceuticals; PamLab, LLC.; Pfizer Inc.; PharmaStar; Pharmavite® LLC; PharmoRx Therapeutics; Precision Human Biology; Prexa Pharmaceuticals, Inc.; Puretech Ventures; PsychoGenics; Psylin Neurosciences, Inc.; RCT Logic, LLC Formerly Clinical Trials Solutions, LLC; Rexahn Pharmaceuticals, Inc.; Ridge Diagnostics, Inc.; Roche; Sanofi-Aventis US LLC; Sepracer Inc.; Servier Laboratories; Schering-Plough Corporation; Solvay Pharmaceuticals, Inc.; Somaxon Pharmaceuticals, Inc.; Somerset Pharmaceuticals, Inc.; Sunovion Pharmaceuticals; Supernus Pharmaceuticals, Inc.; Synthelabo; Taisho Pharmaceutical; Takeda Pharmaceutical Company Limited; Tal Medical, Inc.; Tetragenex Pharmaceuticals, Inc.; Transform Pharmaceuticals, Inc.; Transcept Pharmaceuticals, Inc.; Vanda Pharmaceuticals, Inc.; VistaGen; he has received speaking or publishing fees from Adamed, Co; Advanced Meeting Partners; American Psychiatric Association; American Society of Clinical Psychopharmacology; AstraZeneca; Belvoir Media Group; Boehringer Ingelheim GmbH; Bristol-Myers Squibb; Cephalon, Inc.; CME Institute/Physicians Postgraduate Press, Inc.; Eli Lilly and Company; Forest Pharmaceuticals, Inc.; GlaxoSmithKline; Imedex, LLC; MGH Psychiatry Academy/Primedia; MGH Psychiatry Academy/Reed Elsevier; Novartis AG; Organon Pharmaceuticals; Pfizer Inc.; PharmaStar; United BioSource, Corp.; Wyeth-Ayerst Laboratories; he has equity holdings in Compellis and PsyBrain, Inc.; he has a patent for Sequential Parallel Comparison Design (SPCD), which are licensed by MGH to Pharmaceutical Product Development, LLC (PPD); and patent application for a combination of Ketamine plus Scopolamine in Major Depressive Disorder (MDD), licensed by MGH to Biohaven; and he receives copyright royalties for the MGH Cognitive & Physical Functioning Questionnaire (CPFQ), Sexual Functioning Inventory (SFI), Antidepressant Treatment Response Questionnaire (ATRQ), Discontinuation-Emergent Signs & Symptoms (DESS), Symptoms of Depression Questionnaire (SDQ), and SAFER; Lippincott, Williams & Wilkins; Wolkers Kluwer; World Scientific Publishing Co. Pte.Ltd.

Dr. Benji Kurian has received research grant support from the following organizations: Targacept, Inc., Pfizer, Inc., Johnson & Johnson, Evotec, Rexahn, Naurex, Forest Pharmaceuticals and the National Institute of Mental Health (NIMH). Mary L. Phillips has received funding from NIMH and the Emmerling-Pittsburgh Foundation.

Dr. Maria Oquendo receives royalties for use of the Columbia Suicide Severity Rating Scale and received financial compensation from Pfizer for the safety evaluation of a clinical facility, unrelated to this study. She is the recipient of a grant from Eli Lilly to support a year's salary for the Lilly Suicide Scholar, Enrique Baca-Garcia, M.D., Ph.D.; she has received unrestricted educational grants and/or lecture fees from Astra-Zeneca, Bristol Myers Squibb, Eli Lilly, Janssen, Otsuka, Pfizer, Sanofi-Aventis, and Shire. Her family owns stock in Bristol Myers Squibb.

Over the past three years, Dr. Diego Pizzagalli has received honoraria/consulting fees from Akili Labs Interactive, BlackThorn Therapeutics, Pfizer, and Posit Science for activities unrelated to this project.

Dr. Rajapillai Pillai, Dr. Chuan Huang, Mr. Andrew LaBella, Dr. Mengru Zhang, Dr. Jie Yang, Dr. Melvin McInnis, Dr. Crystal Cooper, Dr. Christine DeLorenzo, and Dr. Ramin Parsey report no relevant or material financial interests that relate to the research described in this paper.

Acknowledgments

We acknowledge the biostatistical computation and support provided by the Biostatistical Consulting Core at School of Medicine, Stony Brook University. We would also like to thank the image analysts at Stony Brook University's Center for Understanding Biology using Imaging Technology (CUBIT) for processing all MRI and DTI data. All Principal Investigators including Madhukar Trivedi, M.D., Patrick McGrath, M.D., Myrna Weissman, M.D., Ramin Parsey, M.D., Ph.D., Maurizio Fava, M.D., were involved with protocol design, data collection, and funding acquisition. The EMBARC study was supported by the National Institute of Mental Health of the National Institutes of Health under award numbers U01MH092221 (Trivedi, M.H.) and U01MH092250 (McGrath, P.J., Parsay, R.V., Weissman, M.M.). Rajapillai Pillai was supported by award number F30MH109412. The funding source had no involvement in research conduct or preparation of this article.

Contributions

Study design: MT, MW, PM, MF, BK, CC, MM, MAO, DAP, RVP. Study conduct: RLIP, CH, AL, CD. Image processing: RLIP, AL, RVP, CD. Statistical analysis: MZ, JY. Data interpretation: RLIP, CH, AL, MZ, JY, RVP, CD. Drafting manuscript: RLIP, CH, AL, CD. RLP takes responsibility for the integrity of the data analysis. All authors read and approved the final manuscript.