A common symptom geometry of mood improvement under sertraline and placebo associated with distinct neural patterns

Lucie Berkovitch, MD, PhD\textsuperscript{a,b,c,d,e}, Kangjoo Lee, PhD\textsuperscript{a,b}, Jie Lisa Ji, PhD\textsuperscript{f}, Markus Helmer, PhD\textsuperscript{f}, Masih Rahmati, PhD\textsuperscript{f}, Jure Demšar, PhD\textsuperscript{g,h}, Aleksij Kraljič\textsuperscript{g}, Andraž Matkovič, PhD\textsuperscript{g}, Zailyn Tamayo, MS\textsuperscript{a,b}, John D. Murray, PhD\textsuperscript{i}, Grega Repovš, PhD\textsuperscript{g}, John H. Krystal, MD, PhD\textsuperscript{a,b}, William J. Martin, PhD\textsuperscript{i}, Clara Fonteneau, PhD\textsuperscript{a,b}, and Alan Anticevic, PhD\textsuperscript{a,b,k}

\textsuperscript{a} Department of Psychiatry, Neuroscience, and Psychology, Yale University School of Medicine, New Haven, CT, USA
\textsuperscript{b} Division of Neurocognition, Neurocomputation, Neurogenetics (N3), Yale University School of Medicine, New Haven, Connecticut, USA
\textsuperscript{c} Université Paris Cité, Paris, France
\textsuperscript{d} Department of Psychiatry, GHU Paris Psychiatrie et Neurosciences, Service Hospitalo-Universitaire, Paris, France
\textsuperscript{e} Unicog, Saclay CEA Centre, Neurospin, Gif-Sur-Yvette Cedex, France
\textsuperscript{f} Manifest Technologies, Inc. New Haven, CT, USA
\textsuperscript{g} Department of Psychology, University of Ljubljana, Ljubljana, Slovenia
\textsuperscript{h} Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia
\textsuperscript{i} Department of Psychological and Brain Science, Dartmouth College, Hanover, NH, USA
\textsuperscript{j} Johnson & Johnson Innovative Medicine, San Diego, CA, USA
\textsuperscript{k} Department of Psychology, Yale University School of Medicine, New Haven, CT, USA

\#: Contributed equally to this study.

Correspondence:

lucie.berkovitch@yale.edu \& fonteneau.clara@gmail.com
40 Temple Street, Suite 6E,
New Haven, CT 06510, USA
(203) 974–7054

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Key Points

Question. What is the data-driven pattern of clinical improvement in major depressive disorder across antidepressants and placebo?

Findings. Clinical improvement has a shared behavioral geometry across groups but differs in terms of intensity between sertraline and placebo. Baseline resting-state global brain connectivity was correlated to improvement scores in the sertraline group only.

Meaning. There is a common behavioral signature of clinical improvement that can be more robustly predicted by baseline neurobehavioral features when it is pharmacologically induced.
Abstract

Importance: Understanding the mechanisms of major depressive disorder (MDD) improvement is a key challenge to determine effective personalized treatments.

Objective: To perform a secondary analysis quantifying neural-to-symptom relationships in MDD as a function of antidepressant treatment.

Design: Double blind randomized controlled trial.

Setting: Multicenter.

Participants: Patients with early onset recurrent depression from the public Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC) study.

Interventions: Either sertraline or placebo during 8 weeks (stage 1), and according to response a second line of treatment for 8 additional weeks (stage 2).

Main Outcomes and Measures: To identify a data-driven pattern of symptom variations during these two stages, we performed a Principal Component Analysis (PCA) on the variations of individual items of four clinical scales measuring depression, anxiety, suicidal ideas and manic-like symptoms, resulting in a univariate measure of clinical improvement. We then investigated how initial clinical and neural factors predicted this measure during stage 1. To do so, we extracted resting-state global brain connectivity (GBC) at baseline at the individual level using a whole-brain functional network parcellation. In turn, we computed a linear model for each brain parcel with individual data-driven clinical improvement scores during stage 1 for each group.

Results: 192 patients (127 women), age 37.7 years old (standard deviation: 13.5), were included. The first PC (PC1) capturing 20% of clinical variation was similar across
treatment groups at stage 1 and stage 2, suggesting a reproducible pattern of symptom improvement. PC1 patients' scores significantly differed according to treatment during stage 1, whereas no difference of response was evidenced between groups with the Clinical Global Impressions (CGI). Baseline GBC correlated to stage 1 PC1 scores in the sertraline, but not in the placebo group.

Conclusions and Relevance: Using data-driven reduction of symptoms scales, we identified a common profile of symptom improvement across placebo and sertraline. However, the neural patterns of baseline that mapped onto symptom improvement distinguished between treatment and placebo. Our results underscore that mapping from data-driven symptom improvement onto neural circuits is vital to detect treatment-responsive neural profiles that may aid in optimal patient selection for future trials.
Introduction

Major depressive disorder (MDD) is a devastating and heterogeneous psychiatric disease.\(^1\) The use of a traditional categorical approach, derived from the DSM-5 criteria\(^2\) severely limits the identification of treatment response predictors.\(^3\) Clinical trials traditionally use predefined scales to compare treatment to placebo. However, an existing gap in our field relates to whether there could be a generalized response pattern that cuts across active treatment and placebo.\(^4,5\) Furthermore, it remains unknown if patients' baseline neural configurations are associated with a unique clinical response pattern. In this study, we intend to test whether a common data driven profile of response can be identified across both treatment and placebo arms. We evaluated data from the publicly available dataset EMBARC (Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care). EMBARC is a multisite randomized placebo-controlled trial in which unmedicated participants with MDD were treated either by sertraline, a serotonin selective reuptake inhibitor (SSRI), or placebo, and then switched to bupropion (an atypical antidepressant) according to clinical response.\(^6\) EMBARC collected clinical measures (including item level data on different scales) and baseline neuroimaging data to find markers associated with antidepressant treatment outcomes.\(^7-13,13-17\) One of the specific aspect of EMBARC is the remarkable intensity of the placebo effect. Using the CGI (Clinical Global Impressions, a clinician rated non-specific 7-point scale providing a single global measure of improvement), or global clinical scores such as the Hamilton Rating Scale for Depression (HRSD), no difference in clinical outcomes was evidenced between the placebo and the sertraline groups.\(^7,9,10,12,13,18\) However, several predicting factors of sertraline and placebo
response were identified, some of them being shared\textsuperscript{10,18} and other being specific to the sertraline or the placebo arm.\textsuperscript{7,9,11,13–16}

Here we applied a new analytic strategy to evaluate whether the placebo and the sertraline arms share a common pattern of symptom improvement. Specifically, we performed a Principal Component Analysis (PCA) on the variations of individual items of four clinical scales, resulting in a univariate measure of clinical improvement. We then compare the sertraline and the placebo arms for this measure and investigated how clinical and neural factors at baseline predicted this measure. With this approach, we were able identify a common clinical profile of symptom improvement that occurs across placebo and sertraline. Critically, however, the baseline neural patterns that were related to symptom improvement strongly distinguished between treatment and placebo.
Material and methods

Data Collection, Study Design and Clinical Sample

In EMBARC, patients with MDD received sertraline or placebo during 8 weeks. Then, their treatment was adapted according to the CGI for 8 additional weeks (see Figure 1). Participants underwent a 3T fMRI at baseline and an extensive clinical assessment at 8 and 16 weeks including the Hamilton Rating Scale for Depression (HRSD), the Altman Self-Rating Mania Scale (ASRM), the Concise Health Risk Tracking (CHRT) and the Concise Associated Symptoms Tracking Scale (CASTS). We included 192 patients who had full clinical data and quality-controlled neuroimaging data at baseline (see Figure 1 and Table 1).
Fig. 1. Study design. In Stage 1, patients with early onset recurrent MDD were randomized to receive either sertraline up to 200 mg daily or placebo under double blind conditions. At week 8, participants were assessed with the CGI and those who had a score of less than “much improved” were considered non-responders. In stage 2, non-responding patients in stage 1 were switched to another treatment under double blind conditions: sertraline non-responders received bupropion, and placebo non-responders received sertraline. Responders continued the treatment received during stage 1. Participants underwent an extensive clinical assessment and a 3T fMRI at baseline. At 8 and 16 weeks, four clinical scales of interest were performed: Hamilton Rating Scale for Depression (HRSD) which assesses depressive symptoms, Altman Self-Rating Mania Scale (ASRM) which measures manic symptoms, Concise Health Risk Tracking (CHRT) which evaluates suicidal propensity and risk and Concise Associated Symptoms Tracking Scale (CASTS) which reflects both anxiety, irritability and manic symptoms.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PLA, N = 99(^1)</th>
<th>SER, N = 93(^1)</th>
<th>p-value(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>(F) 63 (64%)</td>
<td>64 (69%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(M) 36 (36%)</td>
<td>29 (31%)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>37.73 (1.30)</td>
<td>37.73 (1.46)</td>
<td>&gt;0.9</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>1 (1.0%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5 (5.1%)</td>
<td>5 (5.4%)</td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>16 (16%)</td>
<td>19 (20%)</td>
<td></td>
</tr>
<tr>
<td>More than one race</td>
<td>4 (4.0%)</td>
<td>6 (6.5%)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>73 (74%)</td>
<td>63 (68%)</td>
<td></td>
</tr>
<tr>
<td>Subject Education (years)</td>
<td>15.39 (0.25)</td>
<td>14.95 (0.28)</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>(Unknown) 0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Major Depressive Disorder - Severity at baseline</td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>High</td>
<td>41 (41%)</td>
<td>41 (44%)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>58 (59%)</td>
<td>52 (56%)</td>
<td></td>
</tr>
<tr>
<td>Major Depressive Disorder - Chronicity at baseline</td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>Chronic</td>
<td>50 (51%)</td>
<td>49 (53%)</td>
<td></td>
</tr>
<tr>
<td>Non chronic</td>
<td>49 (49%)</td>
<td>44 (47%)</td>
<td></td>
</tr>
<tr>
<td>HRSD - Hamilton Rating Scale for Depression (total score)</td>
<td>18.42 (0.44)</td>
<td>18.73 (0.44)</td>
<td>0.6</td>
</tr>
<tr>
<td>ASRM - Altman Self-Rating Mania Scale (total score)</td>
<td>1.30 (0.18)</td>
<td>1.42 (0.18)</td>
<td>0.6</td>
</tr>
<tr>
<td>CHRT - Concise Health Risk Tracking Scale (Propensity score)</td>
<td>26.96 (0.84)</td>
<td>26.88 (0.81)</td>
<td>&gt;0.9</td>
</tr>
<tr>
<td>CHRT - Concise Health Risk Tracking Scale (Risk score)</td>
<td>5.39 (0.23)</td>
<td>5.54 (0.26)</td>
<td>0.7</td>
</tr>
<tr>
<td>CASTS - Concise Associated Symptoms Tracking Scale (total score)</td>
<td>29.69 (0.91)</td>
<td>30.39 (0.99)</td>
<td>0.6</td>
</tr>
</tbody>
</table>

\(^1\)n (%); Mean (sem)

\(^2\)Pearson’s Chi-squared test; Welch Two Sample t-test

Table 1. Baseline patients characteristics.
Dimension reduction of symptom improvement

To estimate a data-driven pattern of symptom improvement, we run a 3-step analysis pipeline on all 73 items across the HRSD, ASRM, CHRT, and CASTS for each individual: (i) dimension reduction of behavioral measures using PCA, separately for each group, (ii) evaluation of PC reproducibility using cross-validation, (iii) comparison of PC loadings across groups to identify shared/distinct axes of symptom improvement. The methodology used for each of these steps is detailed in the Supplementary. PCA can reveal differences between groups in two ways. First, groups can differ in terms of PC geometry (i.e. loadings), meaning that the type of symptoms that change in time depends on the group. Alternatively, groups can have the same PC geometry (i.e. their PC loadings are highly correlated). In such cases, we can rerun the PCA pooling the participants together and show that they have different score distributions for the resulting common PCs (i.e. different “intensity” of change on a common axis of clinical variation), using post-hoc two-sample two-sided t-tests.

Dimension reduction of neural data

Neuroimaging data were acquired at baseline during resting state using 3T fMRI. All scans were processed using the Quantitative Neuroimaging Environment & ToolboX (QuNex), which integrates the Human Connectome Project Pipelines. The preprocessing pipeline is described in the Supplementary. Neural data was examined at different levels: parcels (718 parcels), networks (12 networks) and whole brain (brain average). Data was first parcellated following the Cole-Anticevic Brain Network Parcellation (CAB-NP) atlas. GBC was then calculated for each subject at the parcel level to reduce the dimensionality of the neural feature space (see Supplementary
about GBC calculation). For network-level analyses, we averaged GBC values across parcels belonging to each network – at the cortical and subcortical levels –, derived from the CAB-NP atlas. For the brain average level analyses, we averaged GBC values across the 718 parcels.

Mass Univariate Behavioral-to-Neural Mapping

The interactions between behavioral scores and individual baseline GBC at the parcel-level were measured via a mass univariate regression procedure, through ANOVA with GBC as a dependent variable, behavioral scores during stage 1 (common PC or response status according to the CGI) as an independent variable and age, gender and site as covariates. The resulting maps correspond to the regression coefficients between patients’ behavioral scores and GBC in each parcel, across all 192 patients. This analysis was also run at the network, the subcortical and brain average levels. The greater the magnitude of the coefficient for a given location, the stronger the statistical relationship between GBC and the behavioral variation across patients during stage 1. Significance of the maps was assessed using PALM. We computed two-tailed Pearson’s correlation, run 1000 permutations and applied family-wise error rate correction.

Results

Analysis of global clinical scores

Differences in global clinical scores at baseline and during each stage are detailed in the Supplementary. There was no difference at baseline between treatment groups. The CHRT was the only scale which improved more in the sertraline than in the
placebo group during stage 1 (delta of CHRT propensity score: \( p = 0.009 \); CHRT risk score: \( p = 0.002 \)). The proportion of responders and non-responders according to the CGI was similar in the two groups at the end of stage 1 (placebo 39.4% vs. sertraline 51.6%, \( \chi^2 = 2.4, p = 0.12 \)). There were no significant differences at baseline between responders and non-responders (all \( p > 0.08 \)), suggesting that baseline symptoms were not predictive of subsequent clinical response.

**Identifying clinical improvement geometry using Principal Component Analysis**

*The first Principal Component has similar loadings across groups*

To further explore the geometry of symptom improvement across groups, we ran separate PCA for each subgroup on the clinical item differences during each stage. Main results are presented in Figure 2 and numeric results are reported in Table 2. Except in patients continuing the same treatment in stage 2, PCA yielded significant and reliable PC1. PC loadings were overall significantly correlated between treatment groups, suggesting that the geometries of symptom improvement were highly similar between groups. To study the distributions of participants’ scores on a common axis of symptom improvement according to their group, we ran an additional PCA on symptom improvement for each stage pooling all the participants together. The common PC1s resulting from this analysis were very similar in terms of loadings to the PC1 observed for each group separately at each stage (see Figure 2, D, I, and Supplementary for a detailed description of items loadings).

*Different PC1 scores distributions between groups*

We then explored whether treatment groups had different scores distributions on common PC1. During stage 1, the sertraline group had on average higher common PC1
scores than the placebo group ($t_{190} = 3.16$, $p = 0.0018$, see Figure 2E). Responders also had higher scores than non-responders (sertraline: responders vs. non-responders: $t_{91} = 8.23$, $p < 0.0001$, placebo: responders vs. non-responders: $t_{80} = 8.57$, $p < 0.0001$, Figure 2, panel G), with a significant difference between sertraline versus placebo non-responders ($t_{97} = 2.25$, $p = 0.027$, whereas sertraline versus placebo responders: $t_{82} = 1.54$, $p = 0.13$). During stage 2, the sertraline group had a higher score than the bupropion group ($t_{80} = 2.39$, $p = 0.019$, Figure 2H). However, since those two groups were defined on previous response, those patients were not comparable in terms of treatment-resistance profiles and sample size. The groups that kept stage 1 treatment did not significantly differ in their score distributions ($t_{80} = 0.82$, $p = 0.42$).
Fig. 2. Symptom improvement during stage 1 and 2: absolute evolution and principal components.

A. Symptom improvement during stage 1 at the item level for each group (red: sertraline, blue: placebo): scores at the item-level are similar between the two treatment groups at baseline and differ only for the CHRT scores evolution. B. Proportion of variance explained by PC representing symptom improvement in each group and split-half cross validation (top: sertraline, bottom: placebo). Only PC1, which explains the most variance, is reliable (r-value > 0.5). C. PC1 loadings: symptom improvement geometry is very similar between the two groups (red: sertraline, blue: placebo). D. PCA pooling all participants yields another PC1 (common PC1: black), which is very similar to the PC1 resulting from PCA run separately on the two groups. E. Distribution of scores for common PC1 in each group (red: sertraline subgroup, blue: placebo).
subgroup, black: all participants). On average, patients under sertraline have higher scores than patients under placebo. F. Results of the same analyses but splitting each treatment group according to clinical response as measured by CGI (salmon: responders to sertraline, brown: non-responders to sertraline, light blue: responders to placebo, dark blue: non-responders to placebo). Responders and non-responders have significant differences of symptom improvement for all clinical scales. PCA again yields a PC1 that explains most variance and is reliable. G. Distribution of scores for the PC1 run across all participants (common PC1) for each subgroup of treatment × response. On average, patients non-responding to sertraline have higher scores than patients non-responding to placebo. H. Results of the same analyses performed during stage 2 (green: patients switched from sertraline to bupropion, orange: patients switched from placebo to sertraline). Patients switched to bupropion have lower global CHRT risk score compared to patients switched to sertraline at the beginning of stage 2 and improved less than participants under sertraline during stage 2. I-J. PCA again yields a PC1 that explains most variance, is reliable and is relatively similar in terms of loadings between the two groups. K. Common PC1 loadings and scores distribution for stage 2 PCA pooling all participants yields another PC1 (common PC1: black), which is very similar to the PC1 resulting from PCA run separately in the different groups. L. Distribution of scores for common PC1 in each group for stage 2 (red: responders to sertraline during stage 1, blue: responders to placebo during stage 2, green: switched to bupropion, orange: switched to sertraline, black: all participants). On average, patients who switched to sertraline have higher scores than patients who switched to bupropion.
<table>
<thead>
<tr>
<th>Group and Stage</th>
<th>Number of significant PCs</th>
<th>Percentage of explained variance</th>
<th>PC1 reliability (split half correlation)</th>
<th>Loadings correlation between PCS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All PC</td>
<td>PC1</td>
<td></td>
</tr>
<tr>
<td><strong>Improvement Stage 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sertraline (n = 93)</td>
<td>6</td>
<td>42.6</td>
<td>19.2</td>
<td>0.87</td>
</tr>
<tr>
<td>Placebo (n = 99)</td>
<td>5</td>
<td>38.1</td>
<td>19.7</td>
<td>0.86</td>
</tr>
<tr>
<td>Sertraline-Responders (n = 48)</td>
<td>3</td>
<td>42.6</td>
<td>14.5</td>
<td>0.56</td>
</tr>
<tr>
<td>Sertraline-Non Responders (n = 45)</td>
<td>6</td>
<td>38.1</td>
<td>19.7</td>
<td>0.60</td>
</tr>
<tr>
<td>Placebo-Responders (n = 39)</td>
<td>5</td>
<td>38.5</td>
<td>15.8</td>
<td>0.56</td>
</tr>
<tr>
<td>Placebo-Non Responders (n = 60)</td>
<td>4</td>
<td>42.7</td>
<td>14.3</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Improvement Stage 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sertraline (previously placebo) (n = 60)</td>
<td>4</td>
<td>43.1</td>
<td>23.5</td>
<td>0.75</td>
</tr>
<tr>
<td>Bupropion (previously sertraline) (n = 45)</td>
<td>4</td>
<td>43.2</td>
<td>19.2</td>
<td>0.56</td>
</tr>
<tr>
<td>Sertraline-Responders (n = 48)</td>
<td>3</td>
<td>28.4</td>
<td>-</td>
<td>&lt; 0.50</td>
</tr>
<tr>
<td>Placebo-Responders (n = 39)</td>
<td>8</td>
<td>66.3</td>
<td>-</td>
<td>&lt; 0.50</td>
</tr>
<tr>
<td><strong>Improvement across all participants (common PC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>7</td>
<td>42.6</td>
<td>19.6</td>
<td>0.93</td>
</tr>
<tr>
<td>Stage 2</td>
<td>7</td>
<td>45.4</td>
<td>20.0</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>Baseline PC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>11</td>
<td>50.1</td>
<td>10.3</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 2. Principal Component Analysis Results.
Baseline predicting factors of clinical improvement

We then explored whether baseline demographic or clinical characteristics could predict PC1 scores during stage 1 by running direct correlation between PC1 scores and baseline clinical scores across subjects and ANOVA with common PC1 scores as a dependent variable and baseline clinical scores and treatment group as independent variables. CHRT risk and CAST global scores had a significant main effect on PC1 scores across groups (CHRT: $r = 0.20$, $p = 0.007$; CAST: $r = 0.20$, $p = 0.006$) with no significant differences between the two groups (interaction CHRT × treatment: $F_{1,188} = 0.6$, $p = 0.45$; interaction CAST × treatment: $F_{1,188} = 0.4$, $p = 0.52$). By contrast HRSD scores and depression severity did not have a main effect on PC1 scores but impacted differently PC1 between the sertraline and the placebo group (interaction HRSD × treatment: $F_{1,188} = 5.8$, $p = 0.017$; interaction depression severity × treatment: $F_{1,188} = 5.6$, $p = 0.019$). They had a higher influence on clinical outcome in the sertraline group (HRSD sertraline group: $r = 0.21$, $p = 0.040$ vs. placebo group: $r = -0.13$, $p = 0.19$; depression severity (high > low) sertraline group: $F_{1,91} = 3.7$, $p = 0.058$ vs. placebo group: $F_{1,91} = 2.1$, $p = 0.15$). So far, we used PCA to explore clinical improvement, with the delta of the symptoms as inputs. However, PCA can also be run on baseline symptoms to explore whether baseline PCs predict subsequent improvement.

Interestingly, the loadings of baseline PC1 and of common improvement PC1 were significantly correlated ($r = 0.69$, $p < 0.001$), indicating that the improvement mainly concerns symptoms that are observable at baseline. Distribution scores for baseline PC1 did not significantly differ between the sertraline and the placebo groups ($t_{186} = -0.82$, $p = 0.41$) but significantly correlated with improvement PC1 scores distribution.
across groups \( (r = 0.20, p = 0.005) \) with no significant differences between the two groups (interaction baseline PC1 scores \( \times \) treatment: \( F_{1,188} = 0.001, p = 0.98 \)). By contrast, clinical response measured by CGI was not significantly predicted by any baseline characteristics (age, gender, ethnicity, education, MDD severity, MDD chronicity, HRSD, ASMR, CHRT, CAST, baseline PC1, all \( p > 0.08 \) across subjects and within each treatment group).

**Brain-behavior mapping**

In order to identify neural predicting factors of response, we then explored the relationships between the baseline parcellated resting-state GBC and clinical improvement. Importantly, we did not find any differences between the sertraline and the placebo groups in baseline GBC (with age, gender and site as covariates) at the parcel, network, and brain average levels (all \( p_{\text{adjusted}} > 0.1 \)).

*Parcellated resting-state GBC*

At the parcel level \( (n = 718) \), the interactions between GBC and common PC1 scores of improvement, on the one hand, or the response status (CGI), on the other hand, did not survive correction for multiple comparisons (family-wise error rate correction, \( \alpha = 0.05 \) in PALM\(^{25} \)). The corresponding neurobehavioral correlation maps are shown in Figure 3A. Additional exploratory analyses revealed a higher correlation between the GBC-PC1 neurobehavioral map and the GBC-response neurobehavioral map in the placebo compared to the sertraline group.

*Network resting-state GBC*
At the network level (averaged GBC across parcels belonging to each network, n = 12 as a dependent variable, PC1 as an independent variable, with age, gender and site as covariates, Figure 3C), there was a significant effect of PC1 scores on GBC in the sertraline group for the auditory ($F_{1,86} = 10.03$, $p_{\text{adjusted}} = 0.026$), the cingulo-opercular ($F_{1,86} = 8.77$, $p_{\text{adjusted}} = 0.045$), and the somatomotor networks ($F_{1,86} = 10.17$, $p_{\text{adjusted}} = 0.024$). In the placebo group, only the secondary visual network GBC significantly interacted with PC1 scores after correction ($F_{1,92} = 13.36$, $p_{\text{adjusted}} = 0.005$). $R$-values were significantly higher in the sertraline than in the placebo group (paired t-test: mean of the difference: 0.06; $t_{11} = 2.97$, $p = 0.013$) with a significant interaction between PC1 scores and treatment for the auditory ($F_{1,183} = 8.84$, $p_{\text{adjusted}} = 0.040$) and the cingulo-opercular networks ($F_{1,183} = 9.02$, $p_{\text{adjusted}} = 0.036$), after Bonferroni correction.

**Subcortical resting-state GBC**

At the subcortical level (averaged GBC across parcels belonging to subcortical structure of interest, i.e. amygdala, hippocampus, thalamus and striatum with its substructures: nucleus accumbens, putamen and caudate nucleus, n = 4, with age, gender and site as covariates, Figure 3D), the only subcortical structure in which GBC significantly interacted with PC1 score was the amygdala in the sertraline group ($F_{1,86} = 7.08$, $p_{\text{adjusted}} = 0.037$). $R$-values were significantly higher in the sertraline than in the placebo group (paired t-test: mean of the difference: 0.10; $t_{10} = 6.70$, $p < 0.001$) with a significant interaction between PC1 scores and treatment for all subcortical regions (all $p_{\text{adjusted}} < 0.008$, after Bonferroni correction).

**Brain average resting-state GBC**
Brain average resting-state GBC (averaged across 718 parcels, with age, gender and site as covariates) significantly predicted clinical improvement in the sertraline group. First, we explore differences of GBC (dependent variable) according to response status (CGI) or PC1 scores (independent variables). GBC was higher for responders compared to non-responders in the sertraline group only (interaction response × treatment across participants: $F_{1,183} = 5.24, p = 0.023$; sertraline group: $F_{1,86} = 7.91, p = 0.006$, placebo: $F_{1,92} = 0.27, p = 0.60$) and was significantly correlated with PC1 scores in the sertraline group only ($r = 0.34, t_{91} = 3.41, p = 0.0010$; placebo: $r = 0.08, t_{97} = 0.75, p = 0.46$). Second, there was a significant effect of GBC on PC1 scores in the sertraline group only (ANOVA with PC1 scores as a dependent variable and GBC, age, gender and site as independent variables: sertraline: $F_{1,86} = 11.42, p = 0.0011$, placebo $F_{1,92} = 0.59, p = 0.44$, Figure 3E), indicating that baseline GBC is predictive of clinical response to pharmacological intervention but not to placebo.
A. Correlation between GBC and PC1 scores

Sertraline

Placebo

Interaction PC1 x treatment

\[ r = 0.56 \]

Correlation between GBC and response status

Sertraline

Placebo

Response x treatment

B. Network parcellation

C. Correlation between GBC and PC1 scores, network level

D. Correlation between GBC and PC1 scores, subcortical level

E. Brain average GBC correlation with PC1 for each subject
Fig. 3. Neurobehavioral mapping of mood improvement during stage 1. A. Top. Correlation between the parcellated resting-state GBC and the PC1 scores in the sertraline (left) and the placebo (middle) groups during stage 1, and interaction between the two treatment groups (right). Bottom. Correlation between the parcellated resting-state GBC and the response status in the sertraline (left) and the placebo (middle) groups, and interaction between the two treatment groups (right). Correlation maps are visually different between the sertraline and the placebo groups, suggesting that the baseline cerebral predictors of clinical improvement differ according to the pharmacological intervention. Additional exploratory analyses showed a lower correlation between GBC-PC1 neurobehavioral mapping and GBC-response neurobehavioral map in the sertraline group ($r = 0.56$, $p < 0.001$) compared to the placebo group ($r = 0.73$, $p < 0.001$), suggesting that response according to CGI has the same brain predictive factors as PC1 for placebo but not for sertraline. The interaction between GBC, PC1 and treatment on the one hand, and GBC, response and treatment on the other hand, displayed in the right panel show how each parcel contributes to the pharmacological response (as opposed to the placebo effect). It should be noted that those interactions did not reach significance after family-wise error rate correction. B. Network parcellation. C. Correlation between the parcellated resting-state GBC regrouped by networks and the PC1 scores in the sertraline (left) and the placebo (right) groups. Each dots represents a parcel and each horizontal bar represents the mean of correlation $r$ values for a given network across subjects, asterisks indicate significant $p$ values (under 0.05), obtained through an ANOVA including also age, gender and site, after Bonferroni correction for multiple comparisons, error bars represent the standard error of the means. There was a significant interaction between PC1 scores and GBC in the sertraline group for the auditory, the cingulo-opercular, and the somatomotor networks, and in the placebo group for one of the two visual network (secondary visual network, Visual 2). $R$ values were significantly higher in the sertraline than in the placebo group when compared with a t-test between the two groups. D. Correlation between the parcellated resting-state GBC regrouped by subcortical regions and the PC1 scores in the sertraline (left) and the placebo (right) groups. Each dots represents a parcel and each horizontal bar represents the mean of correlation $r$ values for a given subcortical region across subjects, asterisks indicate significant $p$ values (under 0.05), obtained through an ANOVA including also age, gender and site, after Bonferroni correction for multiple comparisons, error bars represent the standard error of the
means. The only subcortical region in which GBC significantly interacted with PC1 score was the amygdala in the sertraline group. \( R \)-values were significantly higher in the sertraline than in the placebo group when compared with a t-test between the two groups. **E.** Brain average GBC correlation with PC1 in the two treatment groups during stage 1 (red: sertraline, blue: placebo). Each dot represent a subject, lines represent the linear regressions, and the shaded areas represent the 95% confidence interval. PC1 scores and GBC being significantly correlated in the sertraline group, but not in the placebo group, indicating that baseline GBC is a predictive factor of pharmacologically-induced clinical improvement.
Discussion

We studied patterns of mood improvement in a cohort of patients with MDD from the EMBARC clinical trial, treated either with antidepressant (sertraline or bupropion) or placebo, using PCA across multiple clinical scales. This data-driven approach yielded a single dimension (PC1) that passed split-half cross-validation testing. This robust PC solution provided a low dimensional pattern of mood improvement and was strikingly similar across both treatment and placebo, suggesting that mood improvement may evolve along a common symptom axis.

One of the specific aspect of EMBARC is the intensity of the reported placebo response. Previous examinations of EMBARC found no differences between sertraline and placebo efficacy, and commonalities in predictive factors between the two groups, supporting the idea that a placebo response is embedded into the antidepressant response. In our analysis, no difference between the two groups and no clinical predictive factors could be isolated using the CGI only. However, by computing a PCA solution, we show a higher efficacy in the sertraline group that was associated with both behavioral and neural predictive factors. Critically, we included a wide range of items across multiple scales, which increased the sensitivity of the PCA-derived measure (see Supplementary for additional discussion). Indeed, suicidal risk and anxiety scores account for a significant part of symptom variation in the PCA solution, especially in the sertraline arm. In this sense, the efficacy of sertraline above placebo would mostly consist in an "antisuicidal" and an "anxiolytic" effect, whereas the pure "antidepressant" response may be attributed to a placebo effect. Interestingly, we observed a dissociation between predictive factors and improvement dimensions. In line
with previous studies, HRSD predicts sertraline efficacy as measured by PC1,\textsuperscript{7,30,31} even if HRSD scores were not more improved in the sertraline compared to the placebo group.\textsuperscript{7,9,10,12,13,18}

In turn, we quantified baseline patterns of data-driven functional connectivity and tested if there may be distinct patterns of baseline GBC that would uniquely map onto magnitude of PC-derived symptom improvement in each arm. Indeed we found that PC1 scores were predicted by an increased resting-state GBC at baseline in the sertraline group, notably in the parcels corresponding to the amygdala and to several cerebral networks. Those results are compatible with previous explorations of EMBARC using different approaches,\textsuperscript{11,15,32} and previous studies on independent datasets.\textsuperscript{33–41} However, contrary to previous findings, there was no significant predictive effect of the DMN\textsuperscript{9,11,15,42,43} or of the rostral anterior cingulate cortex connectivity\textsuperscript{13,18,44–48} on clinical improvement. Similarly, previous work shows that a decreased amygdala connectivity at baseline correlates with positive clinical outcomes,\textsuperscript{40,49–52} whereas we, and others,\textsuperscript{53} found an opposite pattern. Those mixed results may be due to the use of GBC for each parcel rather than predetermined ROI or seed, or connectivity measures (within or intrinsic) restricted to networks. More importantly, and in line with previous results,\textsuperscript{8,9,11,13,15} our study shows that baseline neural patterns that relate to sertraline and placebo effects are in fact different. In other words, even if the pattern of the mood improvement is highly similar in the two groups, it seems to be related to very different baseline neural system configuration. The parcel-level analyses further suggest that improvement as measured by CGI and PC1 correspond to distinct neural circuits in the sertraline group but not in the placebo.
Sertraline could therefore activate additional brain mechanisms that may complement or amplify the placebo response (see Supplementary for additional discussion).

Our study has several limitations. First, it is a secondary analysis of a publicly available dataset. This analysis was therefore not initially planned and could have been biased by previously published data on the same cohort. Second, we did not reproduce all previously published findings, probably because we used a different analytical approach and selected a subset of participants who had clinical measures at different time-points, which raises the question of generalizability. In this perspective, our results would highly benefit from a replication with an independent dataset, notably to explore whether the pattern of mood improvement that we observed is specific to this population or more universal. Third, the design did not include neuroimaging at the end of stage 1, so we could not map clinical changes to neural evolution and could only study predictive factors of clinical response. A longitudinal study would allow measuring how neural patterns evolve with time according to treatment and response. Finally, only few conclusions could be drawn from the second stage since participants were assigned to a new treatment according to their clinical response in stage 1, rendering the groups incomparable. Comparing different drugs in randomized parallel arms would allow examining response and predictive factors of various antidepressants, and better characterize their pharmacological mechanisms of action.

In summary, we discovered a common behavioral signature of clinical improvement along the mood spectrum with multiple symptomatic dimensions, on which patients score differently according to the treatment received. At the behavioral level,
improvement under antidepressant therefore corresponds to an amplification of the
placebo response. This improvement was more robustly predicted by baseline GBC (at
different levels) and clinical features in the sertraline group, suggesting that
pharmacological improvement relies on more reproducible and specific
neurobehavioral features.
Declarations

A.A. and J.D.M. hold equity with Neumora Therapeutics (formerly BlackThorn Therapeutics), Manifest Technologies, and are co-inventors on the following patents:

Clearmind Medicine, Cybin IRL, Delix Therapeutics, Eisai, Enveric Biosciences, 
Epiodyne, EpiVario, Evidera, Freedom Biosciences, Janssen Research & Development, 
Jazz Pharmaceuticals, Leal Therapeutics, Neumora Therapeutics, Neurocrine 
Biosciences, Novartis Pharmaceuticals Corporation, Otsuka America Pharmaceutical, 
Perception Neuroscience, Praxis Precision Medicines, PsychoGenics, Spring Care, 
Sunovion Pharmaceuticals, Takeda Industries, Tempero Bio, Terran Biosciences, and 
Tetricus. All other co-authors declare no competing interests.

Acknowledgements

L.B. was supported by the Fondation Bettencourt Schueller, the Philippe Foundation 
and the L’Oréal-Unesco Foundation. G.R., J.D. are funded by ARIS grants P3-0338 and 
J5-4590, A.K. and A.M. by ARIS grants J5-4590 and P5-0110, K.L. is funded by 
1U01MH124639-03 (NIMH/NIH/DHHS); Boehringer Ingelheim International GmbH - 
Project No.668012.

References

1. World Health Organization. Depression and Other Common Mental Disorders: Global Health 
   https://apps.who.int/iris/handle/10665/254610

2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th 


4. Gueorguieva R, Mallinckrodt C, Krystal JH. Trajectories of depression severity in clinical trials of 
   duloxetine: insights into antidepressant and placebo responses. Arch Gen Psychiatry. 


Randomized
N = 192

Sertraline
N = 93

Placebo
N = 99

Stage 1: 8 weeks

Responders
N = 48
(52%)

Non responders
N = 45
(48%)

Responders
N = 39
(39%)

Non responders
N = 60
(61%)

Continue Sertraline

Switch to Bupropion

Continue Placebo

Switch to Sertraline

Stage 2: 8 weeks

Neuroimaging
[T1w & BOLD]

Demographics
HRSD, ASRM, CHRT, CAST
A. Correlation between GBC and PC1 scores

Interaction PC1 × treatment

B. Network parcellation

C. Correlation between GBC and PC1 scores, network level

D. Correlation between GBC and PC1 scores, subcortical level

E. Brain average GBC correlation with PC1 for each subject

$r = 0.56$

$r = 0.73$

$r = 0.34$
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PLA, N = 99</th>
<th>SER, N = 93</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td><em>F</em></td>
<td>63 (64%)</td>
<td>64 (69%)</td>
<td></td>
</tr>
<tr>
<td><em>M</em></td>
<td>36 (36%)</td>
<td>29 (31%)</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>37.73 (1.30)</td>
<td>37.73 (1.46)</td>
<td>&gt;0.9</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>1 (1.0%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5 (5.1%)</td>
<td>5 (5.4%)</td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>16 (16%)</td>
<td>19 (20%)</td>
<td></td>
</tr>
<tr>
<td>More than one race</td>
<td>4 (4.0%)</td>
<td>6 (6.5%)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>73 (74%)</td>
<td>63 (68%)</td>
<td></td>
</tr>
<tr>
<td><strong>Subject Education (years)</strong></td>
<td>15.39 (0.25)</td>
<td>14.95 (0.28)</td>
<td>0.2</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Major Depressive Disorder - Severity at baseline</strong></td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td><em>High</em></td>
<td>41 (41%)</td>
<td>41 (44%)</td>
<td></td>
</tr>
<tr>
<td><em>Low</em></td>
<td>58 (59%)</td>
<td>52 (56%)</td>
<td></td>
</tr>
<tr>
<td><strong>Major Depressive Disorder - Chronicity at baseline</strong></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td><em>Chronic</em></td>
<td>50 (51%)</td>
<td>49 (53%)</td>
<td></td>
</tr>
<tr>
<td><em>Non chronic</em></td>
<td>49 (49%)</td>
<td>44 (47%)</td>
<td></td>
</tr>
<tr>
<td><strong>HRSD - Hamilton Rating Scale for Depression (total score)</strong></td>
<td>18.42 (0.44)</td>
<td>18.73 (0.44)</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>ASRM - Altman Self-Rating Mania Scale (total score)</strong></td>
<td>1.30 (0.18)</td>
<td>1.42 (0.18)</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>CHRT - Concise Health Risk Tracking Scale (Propensity score)</strong></td>
<td>26.96 (0.84)</td>
<td>26.88 (0.81)</td>
<td>&gt;0.9</td>
</tr>
<tr>
<td><strong>CHRT - Concise Health Risk Tracking Scale (Risk score)</strong></td>
<td>5.39 (0.23)</td>
<td>5.54 (0.26)</td>
<td>0.7</td>
</tr>
<tr>
<td><strong>CASTS - Concise Associated Symptoms Tracking Scale (total score)</strong></td>
<td>29.69 (0.91)</td>
<td>30.39 (0.99)</td>
<td>0.6</td>
</tr>
</tbody>
</table>

1n (%); Mean (sem)

2Pearson's Chi-squared test; Welch Two Sample t-test
<table>
<thead>
<tr>
<th>Group and Stage</th>
<th>Number of significant PCs</th>
<th>Percentage of explained variance</th>
<th>PC1 reliability (split half correlation)</th>
<th>Loadings correlation between PCS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All PC</td>
<td>PC1</td>
<td></td>
</tr>
<tr>
<td><strong>Improvement Stage 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sertraline (n = 93)</td>
<td>6</td>
<td>42.6</td>
<td>19.2</td>
<td>0.87</td>
</tr>
<tr>
<td>Placebo (n = 99)</td>
<td>5</td>
<td>38.1</td>
<td>19.7</td>
<td>0.86</td>
</tr>
<tr>
<td>Sertraline-Responders (n = 48)</td>
<td>3</td>
<td>42.6</td>
<td>14.5</td>
<td>0.56</td>
</tr>
<tr>
<td>Sertraline-Non Responders (n = 45)</td>
<td>6</td>
<td>38.1</td>
<td>19.7</td>
<td>0.60</td>
</tr>
<tr>
<td>Placebo-Responders (n = 39)</td>
<td>5</td>
<td>38.5</td>
<td>15.8</td>
<td>0.56</td>
</tr>
<tr>
<td>Placebo-Non Responders (n = 60)</td>
<td>4</td>
<td>42.7</td>
<td>14.3</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Improvement Stage 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sertraline (previously placebo) (n = 60)</td>
<td>4</td>
<td>43.1</td>
<td>23.5</td>
<td>0.75</td>
</tr>
<tr>
<td>Bupropion (previously sertraline) (n = 45)</td>
<td>4</td>
<td>43.2</td>
<td>19.2</td>
<td>0.56</td>
</tr>
<tr>
<td>Sertraline-Responders (n = 48)</td>
<td>3</td>
<td>28.4</td>
<td>-</td>
<td>&lt; 0.50</td>
</tr>
<tr>
<td>Placebo-Responders (n = 39)</td>
<td>8</td>
<td>66.3</td>
<td>-</td>
<td>&lt; 0.50</td>
</tr>
<tr>
<td><strong>Improvement across all participants (common PC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>7</td>
<td>42.6</td>
<td>19.6</td>
<td>0.93</td>
</tr>
<tr>
<td>Stage 2</td>
<td>7</td>
<td>45.4</td>
<td>20.0</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>Baseline PC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>11</td>
<td>50.1</td>
<td>10.3</td>
<td>0.60</td>
</tr>
</tbody>
</table>