Structural abnormalities associated with poor outcome of a major depressive episode: the role of thalamus.
Jean-Marie Batail, Julie Coloigner, Marine Soulas, Gabriel Robert, Christian Barillot, Dominique Drapier

To cite this version:

HAL Id: hal-02920551
https://hal.science/hal-02920551
Submitted on 18 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Structural abnormalities associated with poor outcome of a major depressive episode: the role of thalamus.

Running title: Thalamus in depressive episode outcome

JM. Batail* a,b,c, @, J. Coloigner* b, M. Soulas a, G. Robert a,h,c, C. Barillot b,#, D. Drapier a,h,c#

a. Centre Hospitalier Guillaume Régnier, Academic Psychiatry Department, F-35703 Rennes, France
b. Univ Rennes, INRIA, CNRS, IRISA, INSERM, Empenn U1228 ERL, F-35042 Rennes, France
c. Univ Rennes, “Comportement et noyaux gris centraux” research unit (EA 4712), F-35000 Rennes, France

* These authors contributed equally to this work.
# These authors contributed equally to this work.

@ Corresponding author :
Jean-Marie Batail
Centre Hospitalier Guillaume Régnier
108 avenue du Général Leclerc
35703 Rennes, France
dfax: +33-2-99-33-39-72
jeanmariebatail@gmail.com
Abstract

An identification of precise biomarkers contributing to poor outcome of a major depressive episode (MDE) has the potential to improve therapeutic strategies by reducing time to symptomatic relief. In a cross-sectional volumetric study with a 6 month clinical follow-up, we performed baseline brain grey matter volume analysis between 2 groups based on illness improvement: 27 MDD patients in the “responder” (R) group (Clinical Global Impression-Improvement (CGI-I) score ≤ 2) and 30 in the “non-responder” (NR) group (CGI-I > 2), using a Voxel Based-Morphometry analysis. NR had significantly smaller Grey Matter (GM) volume in the bilateral thalami, in precentral gyrus, middle temporal gyrus, precuneus and middle cingulum compared to R at baseline. Additionally, they exhibited significant greater GM volume increase in the left anterior lobe of cerebellum and posterior cingulate cortex. The latter result was not significant when participants with bipolar disorder were excluded from the analysis. NR group had higher baseline anxiety scores. Our study has pointed out the role of thalamus in prognosis of MDE. These findings highlight the involvement of emotion regulation in the outcome of MDE. The present study provides a step towards the understanding of neurobiological processes of treatment resistant depression.

Keywords: Mood disorders, Major Depressive Episode, Treatment-Resistance, Neuroimaging, Grey Matter, Thalamus.
1. Introduction

Mood Depressive Disorder (MDD) is a very common mental disorder with a lifetime prevalence of 16% (Kessler et al., 2003) and is the leading cause of disability worldwide (Vasic et al., 2008). Only 50% of patients respond to their first treatment and remission rate with standard antidepressant treatments is only 30-40% (Murray et al., 2012; Rush et al., 2006). Reliable predictors of poor outcome in patient with a Major Depressive Episode (MDE) are lacking (Nathan and Gorman, 2007). A biomarker has been defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” (“Biomarkers and surrogate endpoints,” 2001). Over the past decade, there is a growing of interest in the development of precision medicine which could help guiding therapeutic strategies using biomarkers-based approaches (Fonseka et al., 2018; Williams, 2017, 2016). This field has been fed by research on biomarkers combining biological and clinical information (Kozak and Cuthbert, 2016; Micoulaud-Franchi et al., 2018).

Since the past decades, the use of neuroimaging techniques has enabled significant advances toward understanding the pathophysiology of mood disorders and identifying potential biomarkers (Fonseka et al., 2018). Structural MRI (sMRI) methods – especially voxel-based morphometry (VBM) – allow non-invasive and accessible investigations of the distribution of grey matter (GM) with automated, repeatable and accurate results (Whitwell, 2009). They provided cumulative evidence of GM abnormalities in emotion and mood regulation key regions in patients with MDD compared to healthy controls (HC) (Lin et al., 2017; Serra-Blasco et al., 2013). Some authors have shown associations between treatment response in depression (TRD) with gray matter (GM) volume. The GM volume changes have been reported including regions involved in cognitive functioning such as deficit in left (Furtado et al., 2012), right (Fu et al., 2013), or both (J. L. Phillips et al., 2015) hippocampus volume (Sämann et al., 2013) and dorsolateral prefrontal cortex (DLPFC) (Fu et al., 2013; Li et al., 2010). Other key regions involved in emotional processing has been found in TRD such as anterior cingulate gyrus (Fujino et al., 2015; Machino et al., 2014), amygdala (Sandu et al., 2017), temporal gyrus (Liu et al., 2012; Ma et al., 2012; Machino et al., 2014; Sämann et al., 2013), insula (Johnston et al., 2015), and cerebellum (Liu et al., 2012; Machino et al., 2014). All these studies used different designs, such as prospective (Abbott et al., 2014; Fujino et al., 2015; Furtado et al., 2012; Korgaonkar et al., 2015; Li et al., 2010; Phillips et al., 2015; Sämann et al., 2013; Schmaal et al., 2015) and retrospective studies (Johnston et al., 2015; Liu et al., 2012; Machino et al., 2014; Sandu et al., 2017), duration of the follow up was also variable from 5 weeks (Sämann et al., 2013) to 5 years (Serra-Blasco et al., 2016), assessing different therapeutic
modalities such as transcranial magnetic stimulation (Furtado et al., 2012), electroconvulsive therapy (ECT) (Abbott et al., 2014), antidepressant (ATD) (Korgaonkar et al., 2015; Li et al., 2010; Sämann et al., 2013), or cognitive behavioral therapy (Fujino et al., 2015), whereas in other patients received personalized therapeutic strategies (Phillips et al., 2015), which could limit the generalizability of the current literature. Thus, these studies did not use consistent definitions of treatment resistance such as resistance to at least two different ATD trial or criteria from Massachusetts General Hospital staging or according to depression rating scales decrease. Therefore, there is a need to transfer scientific knowledge on biomarkers in treatment resistant depression over routine care clinics applied to a Major depressive Episode (MDE).

In this paper, we performed a cross-sectional volumetric study with a 6-month clinical follow-up to examine the relationship between voxel-wise GM volume and a poor outcome of depression. The sample of patients were treated with personalized care according to therapeutic guidelines and were followed over 6 months, the usual timescale of one MDE (Solomon, 1997; ten Have et al., 2017). In this study, we compared whole brain volume differences at baseline between 2 groups of depressed patients – responders (R) / non-responders (NR) – based on illness improvement at 6-month follow-up. Regarding to the current literature on treatment resistance in Mood Depressive Disorder, we hypothesized that patients who do not achieve remission from, will have different grey matter volume at baseline, affecting key regions involved in emotional and cognitive processes.

2. Methods

2.1. Patient population

Patients included suffered from an MDE according to DSM-5 criteria with or without a personal history of depressive mood disorder (unipolar or bipolar subtype). Exclusion criteria included other Axis-I disorders (except for anxious comorbidities such as post-traumatic stress disorder, social phobia, generalized anxiety disorder or panic disorder), which were explored using the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Patients with severe chronic physical illness were not included. Other exclusion criteria were potential safety contraindications for MRI (pacemakers, metal implants, pregnancy and lactation), diagnosed neurodegenerative disorders (e.g. Parkinson’s disease, Alzheimer’s disease, Huntington’s disease), a history of significant head injury (history of head injury with loss of consciousness which required hospitalization), or diagnosed dementia (according to DSM-5 criteria).
2.2. Study design

Depressed patients were recruited from routine care units in the psychiatric university hospital in Rennes between November 2014 and January 2017 and were enrolled in a prospective open cohort study. A complete description of the study has been given to the subjects, and their written informed consent was obtained. The study was approved by an ethic committee (Comité de Protection des Personnes Ouest VI – ref : CCP Ouest 6 – 829) and is registered in www.clinicaltrial.gov (NCT02286024). When recruited, patients underwent a first structured clinical interview at baseline. By a maximum of three days after clinical assessment patients had an imaging protocol. They received personalized care according to therapeutic guidelines described elsewhere (Phillips et al., 2015). Patients were followed-up in routine care by their physicians (as frequently as required by their clinical state). If they didn’t meet 20% improvement at MADRS scale, dose was either increased, or medication changed. Medication changes were based on international guidelines (Bauer et al., 2017; Bennabi et al., 2019; Charpeaud et al., 2017; Kennedy et al., 2016) and NICE guidelines1. Following these recommendations, monotherapy or combination/association therapies were then chosen according to their treatment resistance level. Before being recruited, both R and NR individuals have been treated according to the same algorithms described above. Therefore, NR have, benefited from adapted therapeutic strategies. Clinical data were anonymously retrieved in a notebook. At 6 months follow-up patients underwent a second structured clinical interview.

2.3. Clinical assessment

Patients were assessed at baseline and at 6 months by a trained clinician (psychiatrist or psychiatry resident). The 6 months' timescale was chosen because it is the median duration of a MDE (ten Have et al., 2017). Sociodemographic (age, gender) and disease characteristics (diagnosis, disease duration, number of mood depressive episodes, actual medication, antecedent of suicidal attempts, treatment resistance stage according to Thase and Rush’s classification) were collected.

In addition, these following scales were assessed:

- Mini-International Neuropsychiatric Interview (M.I.N.I.) (Sheehan et al., 1998)
- Montgomery and Åsberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979) and Beck Depression Inventory (BDI) (Beck et al., 1961)
- Widlöcher Depressive Retardation Scale (WDRS) (Widlöcher, 1983)

1 https://www.nice.org.uk/guidance/indevelopment/gid-cgwave0725/documents
- State Trait Anxiety Inventory A and B (STAI) (Spielberger, C. D. et al., 1983) and Hamilton Anxiety Scale (HAMA) (Hamilton M., 1959)
- Snaith Hamilton Pleasure Scale (SHAPS) (Snaith et al., 1995)
- Apathy Evaluation Scale (AES) (Marin et al., 1991)
- Clinical Global Impression – Severity (CGI-S) and the Clinical Global Impression - Improvement (CGI-I), (Guy, 1976).

The Clinical Global Impression – improvement subscale (CGI-I) scale was used after a 6-months period to measure the treatment response. The CGI-I measure is rated from 1 (very much improved) to 7 (very much worse). This measure is indeed a well-established research brief rating tool applicable to all psychiatric disorders and well correlated with other standard scales (Busner and Targum, 2007). It provides a global rating of illness severity and improvement, taking into accounts the patient’s history, social circumstances, symptoms and the impact of the illness on the patient’s ability to function (Busner and Targum, 2007). Leucht et al. also demonstrated that a CGI-I score of 2 (much improved) corresponded to a 50-60% reduction in the HAMD-17 score, while a score of 1 reflected a decrease from 75 to 85 % of the same scale within the 6-months longitudinal follow-up of depressive patients (Leucht et al., 2013). Min et al. also used this threshold of a score ≤ 2 after a 6 months' follow-up to define a favorable clinical course and a therapeutic response (Min et al., 2012).

2.4. Imaging procedure

2.4.1. Data acquisition

Patients were scanned on a 3T whole body Siemens MR scanner (Magnetom Verio, Siemens Healthcare, Erlangen, Germany) with a 32-channel head coil. Anatomical data included a high-resolution 3D T1-weighted MPRAGE sequence (3D T1w) with the following imaging parameters: TR/TE/TI = 1900/2.26/900 ms, 256x256 mm² FOV and 176 sagittal slices, 1x1x1 mm³ resolution, parallel imaging GRAPPA2.

2.4.2. Data processing

Imaging data were analyzed by optimized voxel-based morphometry (VBM), using Statistical Parametric Mapping software (SPM) (http://www.fil.ion.ucl.ac.uk/spm/) implemented in Matlab (MathWorks, Natick, Massachusetts). The MRI scans were segmented into GM, white matter and cerebrospinal fluid using a unified tissue-segmentation module, after correcting for image-intensity non-uniformity. An Exponentiated Lie algebra (DARTEL) template was then
computed from the GM and white matter of the R and NR patients (Ashburner, 2007). Then, a
two-step registration method was used to transform segmented images into the Montreal
Neurological Institute (MNI) template space. First, segmented images of each subject were
then spatially non-linearly normalized to customized template; then affine transformation
between this DARTEL template and Montreal National Institute (MNI) space was calculated.
To preserve gray matter volume within each voxel, we modulated images by Jacobian
determinants derived from the spatial normalization by DARTEL and then smoothed them by
using an 8-mm FWHM Gaussian kernel.

2.5. Statistical analyses

Statistical analyses were performed with R software (http://www.R-project.org/) only on
patients who had undergone both baseline and 6 months assessments. All results are reported
as means ± standard derivation (SD) for continuous variables, and rates for discrete variables.
The significance threshold for all tests was set at 5%.

2.5.1. Statistical analysis of clinical data

The whole sample was divided into two groups using a 2-CGI-I cut-off score at the 6 months
assessment, as suggested by Min et al. (Min et al., 2012) (CGI-I ≤ 2 : R group; CGI-I > 2 : NR
group). Socio-demographic and clinical variables were then compared between these 2
groups.

Shapiro Wilk test was used for each variable to determine if the data set is well-modeled by a
normal distribution and the QQ plots were visually inspected to ensure normal distribution.
Quantitative variables were then compared using either the Student’s test or the Wilcoxon test,
when needed. Owing to the different sizes of the groups, qualitative variables were compared
with either the Chi-squared test or a Fisher test, when needed.

2.5.2. Morphometric data analysis

A whole-brain voxel-wise analysis was performed in SPM to detect baseline differences in GM
between the NR and R groups with the 2- CGI-I cut-off score. Age, gender, total intracranial
volume (TIV), and medication load were included as covariates in each comparison. A
composite measure of medication load for each patient was assessed using a previously
established method (Sackeim, 2001). The two-sample t-tests were conducted within a group
GM mask obtained by selecting a threshold of 0.2 on the mean GM map of all subjects.
Thresholds were set at a corrected $p < 0.01$, corrected for multiple comparison using the AlphaSim program in AFNI (Analysis of Functional NeuroImages), determined by Monte Carlo simulation (Parameters were: single voxel $p$-value = 0.01, a minimum cluster size of 7.3 mm$^3$, FWHM = 8 mm, within a gray matter mask corresponding to the MNI atlas). The student $t$-value reported in the result section corresponds to the maximum value of each cluster. To explore to what extent these results are associated with individuals with bipolar disorders, we performed the same analysis after removing patients suffering from bipolar disorder.

3. Results

3.1 Demographics and clinical measures

Initially, 117 patients were included at baseline. Of 53 patients were lost to follow-up, 3 died during the timeframe of the study: 2 by suicide and 1 for non-psychiatric reason. Forty-three patients could not be traced, most commonly because they were not routinely followed in the academic hospital. Sixty-four patients were present at follow up. Of these, 2 did not have an imaging protocol and CGI-I score was missing for 5 of them. In total, 57 patients were analyzed. According to the 2-CGI-I cut off score, 27 were R patients and 30 NR ones. There were no significant differences between patients analyzed and those lost at follow-up regarding the baseline depression severity (MADRS total scores) (two sample t.test, $t = 1.55$, df = 107.72, $p$-value = 0.12, CI95 [-0.46; 3.77]).

The study’s flowchart is shown on Figure 1. Socio-demographical and clinical characteristics at baseline and at 6 months of all analyzed patients are summarized in Table 1. The sample was predominantly represented by women, middle-aged and suffering from moderate depression.

Baseline demographic data and clinical characteristics from 27 R patients and 30 NR patients were compared. Results are shown in Table 2. The NR group was significantly younger in comparison with the R group. No significant difference was found between the two groups on gender, education level, age of onset, comorbidities, medication load, illness duration and number of depressive episodes. In addition, severity of depression (measured by MADRS and BDI scores) was not significantly different between the R group and NR group. NR patients were significantly more anxious at baseline according to the Hamilton Anxiety scale (HAM-A).
3.2. Morphometric results: group comparisons of GM volume using VBM

Group comparison was controlled for age, gender, TIV, and medication load. The detailed t-test results, including the coordinates of the clusters are presented in table 3 and in figures 2 and 3.

In comparison with R patients at baseline, NR revealed a significant reduction of GM volume in an extended network including bilateral thalami, right frontal lobe (precentral gyrus), parietal lobes (precuneus and mid cingulum), and temporal lobe (mid temporal gyrus). In addition, with NR patients showed greater GM volume cluster in left anterior cerebellum and left posterior cingulate cortex. This latter result was not significant when participants with bipolar disorder were excluded from the analysis.

4. Discussion

4.1. General considerations

In addition to prior knowledge on that topic, we found significant baseline GM volume reduction in bilateral thalamus and an increase in cerebellum at baseline in patients suffering from a MDE with a poor outcome at 6 months. Our results appeared in line with our hypothesis and previous findings from structural neuroimaging studies (Chen et al., 2007). Altogether, there is converging evidence that poor outcome of depression is associated with gray matter deviations affecting key regions involved in emotional and cognitive processes.

4.2. Cortical gray matter volume reductions

In the NR group, we observed reduced GM volume over the frontal (precentral gyrus), the temporal (middle temporal gyrus), and the parietal lobes (precuneus and middle cingulum) at baseline. These results were consistent with Peng et al who reported that the frontal-limbic circuit abnormality is implicated in the pathogenesis of MDD and are closely associated with clinical manifestations, including emotional bias, rumination and cognitive deficit (Peng et al., 2011). A recent meta-analysis of VBM studies in adults with MDD (Fu et al., 2013) revealed an association between a decrease in GM volume in the left DLPFC and reduced likelihood of antidepressant response (Chen et al., 2007; Costafreda et al., 2009; Li et al., 2010). Another structural MRI study using machine learning approach also reported association between reduced GM volume in right DLPFC and poor response after a 6 weeks period of treatment (Gong et al., 2011). According to Li et al., structural deficits in the left DLPFC might predict
poor or delayed antidepressant responses in adult patients with recurrent MDD (Li et al., 2010). Furthermore, the same authors have highlighted the volume variation of some small regions within parietal lobe (postcentral gyrus) correlated with depressive symptoms or visual/acoustic attention (Li et al., 2010).

4.3. Thalamic volume reduction

The thalamus is a pair mass of gray matter in the dorsal part of the diencephalon which has mainly a function of relay and integration of sensory and motor signals to the cerebral cortex and a function of regulation of consciousness, sleep, and alertness (Taber et al., 2004). It is also a key structure in memory and emotion and is known to be involved in the pathophysiology of mood disorders (Taber et al., 2004). In early 1980s, Angelini et al. characterized a depressive syndrome as a very frequent side effect after stereotaxic thalamotomy in patients with abnormal movements (Angelini et al., 1982). Later, Young et al. pointed out the role of the anterior and mediodorsal thalamus in the expression and experience of emotion (Young et al., 2004).

Our findings exhibited a GM volume reduction in ventral lateral nucleus and medial dorsal nucleus. In contrast with the large number of studies (Bora et al., 2012), only two recent analyses have revealed volumetric reductions in this regions in MDD (Nugent et al., 2013; Webb et al., 2014). Besides, a recent study suggested a significant correlation between severity of depression and anxiety symptoms and thalamic reductions in GM volume in an adolescent MDD group compared to healthy controls (Hagan et al., 2015). This study is in accordance with our clinical results in which NR patients had significant higher anxiety score on the HAM-A scale at baseline. This result suggests that thalamic volume reduction plays a role in the outcome of a MDE possibly via its involvement in the pathophysiology of emotion such as anxiety. It could be interesting to explore the proper role of the thalamus in anxious dimension and prediction of evolution of one episode.

4.4. The role of cerebellum in depression

After correcting for bipolar comorbidity, GM volume increase in cerebellum in the NR group at baseline did not remain significant. This result can be discussed regarding two aspects, either the loss of power of our statistical analysis or the effect of bipolar disorder (BD) on GM volume in depressed patients. Therefore, relative increasing of GM volume between NR and R patient could be almost explained by the bipolar condition. In line with this, a recent meta-analysis has shown that GM volume decrease in cerebellum is more pronounced in MDD than BD, relative
to healthy controls (Wise et al., 2017). As we didn't have a control group, BD could explain the relative increase in cerebellum in NR patients. In contrast to the large number of imaging studies exploring cerebral cortex in MDD, fewer morphometric MRI studies have assessed cerebellum in depression. Given its well-described role in motor control; particularly in coordination, precision, and accurate timing, cerebellum was recently well-recognized to participate in the organization of higher order functions; in both cognitive and affective processing control (Adamaszek et al., 2017; J. R. Phillips et al., 2015). More precisely, prior studies showed cerebellar involvement in elaborating negative emotions (Turner et al., 2007; Xu et al., 2017). Impairments in subjective experiences of pleasant emotions in response to positive stimuli have been reported in patients with cerebellar lesions (Adamaszek et al., 2017; Turner et al., 2007). According to Clausi et al., this region has a role in modulating the unconscious and conscious levels of emotional processing (Clausi et al., 2017). However, previous whole brain structural MRI studies investigating GM volumes abnormalities in MDD have reported different results regarding the cerebellum ranging from atrophy (Fossati, 2015; T. S. Frodl et al., 2008; Peng et al., 2011) to no significant difference (Guo et al., 2013) in MDD patients compared to healthy controls. Then, Depping et al. showed that cerebellar volume increase following ECT was associated with HAM-D score reduction (Depping et al., 2017). Taken together, accumulating evidences are converging toward volume abnormalities located in cerebellum in depression. However, our study, in line with Wise et al., highlighted that it may to be more related to bipolar depression than MDD disorder (Wise et al., 2017).

4.5. Limitations

Our results have to be discussed through some limitations. First, the relative small sample-size may have affected the statistical power of our analyses. Second, anxiety levels should be taken into consideration when interpreting our results. Third, patients recruited were not drug-free but all were treated with antidepressants, antipsychotics or mood stabilizers. Nevertheless, this has been taken into account in our analyses. Furthermore, it should be noted that despite patients benefited from adapted therapeutic strategies, this have not been systematically controlled because of the observational design. Fourth, the absence of healthy control group limits the generalizability of these findings. Fifth, we have to notice that according to our CGI-I criteria of treatment response, our R group meets for 73% the criteria of remission. Therefore, our findings could be viewed as biomarkers of remission rather than treatment response.

5. Conclusion
This cross-sectional study focused on structural abnormalities associated with poor outcome at 6 months in patients with MDE. The most striking result were that NR patients were characterized by GM volume decrease in thalamus. Thus, patients were also more anxious than R patients at baseline. Consistent with our clinical results and prior findings, the implication of these regions in outcome of an MDE could be linked with the severity of the depressive state through their involvement in anxious dimension. In addition to their role in the physiopathology of depression, our results have highlighted the implication of thalamic-fronto-cortico-cerebellar circuit in the prognosis of an episode. Future studies could go further by focusing on anxiety dimension in order to identify the proper role of the thalamus and cerebellum in a sample of greater size to unambiguously rule out the lack of power. These findings suggest that GM abnormalities are of high interest for the research of "routine care biomarkers based-approaches" of poor outcome of an MDE. These biomarkers could be very useful in order to define biotypes of patients at risk of treatment resistance.

Acknowledgments

MRI data acquisition was supported by the Neurinfo MRI research facility from the University of Rennes I. Neurinfo is granted by the European Union (FEDER), the French State, the Brittany Council, Rennes Metropole, Inria, Inserm and the University Hospital of Rennes. This work has been funded by Institut des Neurosciences Cliniques de Rennes (INCR). The authors thank Mr Stéphane Brousse and Mr Jacques Soulabaille for their involvement in the conduct of the study. The authors have no competing interests to declare. This work is dedicated to the memory of Christian Barillot, director of the former Visages now Empenn Imaging research unit in Rennes from 2006 to 2020.

References


<table>
<thead>
<tr>
<th>Variables (n=57)</th>
<th>M0</th>
<th>M6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sociodemographic</strong></td>
<td>Mean/n</td>
<td>sd / %</td>
</tr>
<tr>
<td>Age (years)</td>
<td>50.95</td>
<td>14.52</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>43</td>
<td>75.44%</td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>16.92</td>
<td>15.37</td>
</tr>
<tr>
<td>Number of episodes</td>
<td>4.56</td>
<td>4.52</td>
</tr>
<tr>
<td>Number of suicidal attempts</td>
<td>1.22</td>
<td>2.11</td>
</tr>
<tr>
<td>Thase and Rush score &gt; 2</td>
<td>24</td>
<td>47.10%</td>
</tr>
<tr>
<td><strong>Comorbidities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder: type 1</td>
<td>3</td>
<td>5.26%</td>
</tr>
<tr>
<td>type 2</td>
<td>10</td>
<td>17.54%</td>
</tr>
<tr>
<td>type 3</td>
<td>7</td>
<td>12.28%</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>8</td>
<td>14.04%</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>23</td>
<td>40.35%</td>
</tr>
<tr>
<td>Social phobia</td>
<td>7</td>
<td>12.28%</td>
</tr>
<tr>
<td>PTSD</td>
<td>5</td>
<td>8.77%</td>
</tr>
<tr>
<td><strong>Medication</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressant: - SSRI</td>
<td>8</td>
<td>14.04%</td>
</tr>
<tr>
<td>- NSRI</td>
<td>22</td>
<td>38.60%</td>
</tr>
<tr>
<td>- TCA</td>
<td>9</td>
<td>15.79%</td>
</tr>
<tr>
<td>- others</td>
<td>13</td>
<td>22.81%</td>
</tr>
<tr>
<td>Mood stabilizer: - Lithium</td>
<td>6</td>
<td>10.53%</td>
</tr>
<tr>
<td>- Anticonvulsant</td>
<td>22</td>
<td>38.60%</td>
</tr>
<tr>
<td>- Antipsychotic</td>
<td>23</td>
<td>40.35%</td>
</tr>
<tr>
<td>Benzodiazepine</td>
<td>30</td>
<td>52.63%</td>
</tr>
<tr>
<td><strong>Clinical variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MADRS</td>
<td>26.67</td>
<td>5.65</td>
</tr>
<tr>
<td>BDI</td>
<td>17.00</td>
<td>7.46</td>
</tr>
<tr>
<td>CGI-S</td>
<td>4.48</td>
<td>2.25</td>
</tr>
<tr>
<td>HAMA</td>
<td>25.49</td>
<td>14.18</td>
</tr>
<tr>
<td>STAI-YB</td>
<td>61.54</td>
<td>11.49</td>
</tr>
<tr>
<td>SHAPS</td>
<td>5.28</td>
<td>3.87</td>
</tr>
<tr>
<td>WDRS</td>
<td>22.07</td>
<td>9.35</td>
</tr>
<tr>
<td>AES</td>
<td>40.65</td>
<td>8.66</td>
</tr>
<tr>
<td>YMRS</td>
<td>1.93</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Table 1: Population description at baseline and at 6 months. sd: standard derivation; PTSD: Posttraumatic stress disorder; SSRI: Selective serotonin reuptake inhibitors; NSRI: Serotonin-norepinephrine reuptake inhibitors; TCA: Tricyclic antidepressants; MADRS: Montgomery-Åsberg Depression Rating Scale; BDI: Beck Depression Inventory; CGI-S: Clinical Global Impression-Severity; HAMA: Hamilton Anxiety Rating Scale; STAI: State-Trait Anxiety Inventory; SHAPS: Snaith Hamilton Pleasure Scale; WDRS: Widløcher Depressive Retardation Scale; AES: Apathy Evaluation Scale; YMRS: Young Maniac Rating Scale
Table 2: Intergroup comparison: demographic and clinical characteristics at baseline between responder and non-responder patients.

<table>
<thead>
<tr>
<th>Variables at baseline (n=57)</th>
<th>Responder (n=27)</th>
<th>Non-responder (n=30)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociodemographic variables</td>
<td>Mean/n</td>
<td>Mean sd % n</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>54.33 15.98 27</td>
<td>47.90 12.57 30</td>
<td>0.049b</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>19 70.37% 27</td>
<td>24 80% 30</td>
<td>0.399c</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.88 4.30 25</td>
<td>11.93 2.12 29</td>
<td>0.559b</td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>20.90 18.60 27</td>
<td>13.33 10.87 30</td>
<td>0.221b</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder type 1</td>
<td>8 29.63% 27</td>
<td>12 41.37% 29</td>
<td>0.359c</td>
</tr>
<tr>
<td>Bipolar disorder type 2</td>
<td>3 11.11% 27</td>
<td>7 24.14% 29</td>
<td>0.299d</td>
</tr>
<tr>
<td>Bipolar disorder type 3</td>
<td>4 14.82% 27</td>
<td>3 10.34% 29</td>
<td>0.700d</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>2 9.09% 22</td>
<td>3 10.71% 28</td>
<td>1d</td>
</tr>
<tr>
<td>Generalized anxiety</td>
<td>11 40.74% 27</td>
<td>12 40.00% 28</td>
<td>0.955c</td>
</tr>
<tr>
<td>Social phobia</td>
<td>3 11.11% 27</td>
<td>4 13.79% 30</td>
<td>1d</td>
</tr>
<tr>
<td>PTSD</td>
<td>1 3.70% 27</td>
<td>4 13.33% 28</td>
<td>0.356d</td>
</tr>
<tr>
<td>Psychotic symptoms</td>
<td>2 7.41% 27</td>
<td>2 6.67% 30</td>
<td>1d</td>
</tr>
<tr>
<td>Medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- SSRI</td>
<td>5 18.52% 27</td>
<td>3 10.34% 29</td>
<td>0.462d</td>
</tr>
<tr>
<td>- NSRI</td>
<td>13 48.15% 27</td>
<td>9 31.03% 29</td>
<td>0.190c</td>
</tr>
<tr>
<td>- TCA</td>
<td>5 18.51% 27</td>
<td>4 13.79% 29</td>
<td>0.725d</td>
</tr>
<tr>
<td>- others</td>
<td>6 22.22% 27</td>
<td>7 24.14% 29</td>
<td>0.865c</td>
</tr>
<tr>
<td>Mood stabilizer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Lithium</td>
<td>4 15.39% 26</td>
<td>2 5.13% 29</td>
<td>0.406d</td>
</tr>
<tr>
<td>- Anticonvulsant</td>
<td>11 40.74% 27</td>
<td>11 37.93% 29</td>
<td>0.830c</td>
</tr>
<tr>
<td>- Antipsychotic</td>
<td>11 40.74% 27</td>
<td>12 41.38% 29</td>
<td>0.961c</td>
</tr>
<tr>
<td>Benzodiazepine</td>
<td>17 62.96% 27</td>
<td>13 44.83% 29</td>
<td>0.174c</td>
</tr>
<tr>
<td>Medication Load</td>
<td>3.41 1.05 27</td>
<td>3.00 1.19 30</td>
<td>0.198b</td>
</tr>
<tr>
<td>Clinical variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MADRS</td>
<td>26.86 5.74 22</td>
<td>26.50 4.94 28</td>
<td>0.815a</td>
</tr>
<tr>
<td>BDI</td>
<td>15.76 7.45 25</td>
<td>18.15 7.43 27</td>
<td>0.253a</td>
</tr>
<tr>
<td>HAMA</td>
<td>21.30 8.89 20</td>
<td>28.14 10.74 22</td>
<td>0.029a</td>
</tr>
<tr>
<td>STAI-YA</td>
<td>56.42 12.09 26</td>
<td>57.14 12.88 29</td>
<td>0.833a</td>
</tr>
<tr>
<td>STAI-YB</td>
<td>59.15 11.25 26</td>
<td>63.60 11.47 30</td>
<td>0.115b</td>
</tr>
<tr>
<td>SHAPS</td>
<td>4.37 3.53 27</td>
<td>6.10 4.03 30</td>
<td>0.100b</td>
</tr>
<tr>
<td>WDRS</td>
<td>24.00 9.84 27</td>
<td>20.33 8.70 30</td>
<td>0.144a</td>
</tr>
<tr>
<td>AES</td>
<td>39.04 6.32 27</td>
<td>42.10 10.21 30</td>
<td>0.175b</td>
</tr>
</tbody>
</table>

Table 2: Intergroup comparison: demographic and clinical characteristics at baseline between responder and non-responder patients. sd: standard derivation; PTSD: Posttraumatic stress disorder; SSRI: Selective serotonin reuptake inhibitors; NSRI: Serotonin–norepinephrine reuptake inhibitors; TCA: Tricyclic antidepressants; MADRS: Montgomery-Åsberg Depression Rating Scale; BDI: Beck Depression Inventory; CGI-S: Clinical Global Impression-Severity; HAMA: Hamilton Anxiety Rating Scale; STAI : State-Trait Anxiety Inventory; SHAPS: Snaith Hamilton Pleasure Scale; WDRS: Widlöcher Depressive Retardation Scale; AES: Apathy Evaluation Scale; a: Student test; b: Wilcoxon test ; c: Chi2 test ; d: Fisher test
Table 3: Regions of statistically significant differences of volume at baseline between responder (R) and non-responder (NR) patients according to VBM analyses. VBM: Voxel-Based Morphometry; L: Left; R: Right; MNI: Montreal Neurological Institute coordinate system or template; Tmax: maximal statistical value of peak voxel showing differences of gray matter volume between two groups.

<table>
<thead>
<tr>
<th>Anatomic regions</th>
<th>Side</th>
<th>Cluster size</th>
<th>Coordinates of peak voxel in MNI space (mm)</th>
<th>Tmax score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td><strong>NR &lt; R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>Ventral Lateral Nucleus</td>
<td>L</td>
<td>300</td>
<td>-19</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>574</td>
<td>9</td>
<td>-14</td>
</tr>
<tr>
<td></td>
<td>Medial Dorsal Nucleus</td>
<td>R</td>
<td>647</td>
<td>9</td>
</tr>
<tr>
<td>Frontal Lobe</td>
<td>Precentral gyrus</td>
<td>R</td>
<td>1995</td>
<td>67</td>
</tr>
<tr>
<td>Temporal Lobe</td>
<td>Middle Temporal Gyrus</td>
<td>R</td>
<td>729</td>
<td>70</td>
</tr>
<tr>
<td>Parietal Lobe</td>
<td>Precuneus</td>
<td>L</td>
<td>1456</td>
<td>-36</td>
</tr>
<tr>
<td></td>
<td>Middle Cingulum</td>
<td>R</td>
<td>493</td>
<td>6</td>
</tr>
<tr>
<td><strong>NR &gt; R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td>Anterior lobe</td>
<td>L</td>
<td>97</td>
<td>-7</td>
</tr>
<tr>
<td>Limbic Lobe</td>
<td>Posterior Cingulate Cortex</td>
<td>L</td>
<td>1206</td>
<td>-5</td>
</tr>
</tbody>
</table>

Table 3: Regions of statistically significant differences of volume at baseline between responder (R) and non-responder (NR) patients according to VBM analyses. VBM: Voxel-Based Morphometry; L: Left; R: Right; MNI: Montreal Neurological Institute coordinate system or template; Tmax: maximal statistical value of peak voxel showing differences of gray matter volume between two groups.
Figure 1: Flow chart

- Patients excluded: $n = 17$
  - 3 neurological comorbidities (1 dementia, 1 Parkinson Disease, 1 epilepsy)
  - 1 English speaker
  - 4 didn’t meet MDE criteria
  - 1 Parkinson’s disease
  - 8 consent refusals

- Patients lost to follow-up at M6: $n = 53$
  - 6 consent withdrawals
  - 3 deaths
  - 1 move
  - 43 could not be traced

- Patients excluded (missing data): $n = 7$
  - 5 without a CGI-I score
  - 2 without MRI acquisition
  - … (anxiety, claustrophobia)

- M0: Patients screened for inclusion
  $n = 134$

- M0: Patients included
  $n = 117$

- M6: Patients assessed for follow-up
  $n = 64$

- M6: Patients analysed for baseline morphometric and clinical data
  $n = 57$

- "Responder" group (CGI-I $\leq 2$)
  $n = 27$

- "Non-responder" group (CGI-I $> 2$)
  $n = 30$
Figure 2: Regions showing significantly decreased GM volume in non-responder (NR) patients compared to responder (R) patients. P: posterior, A: anterior, L: left, R: right. T: thalamus, Pr: precuneus, PCG: posterior central gyrus, MTG: middle temporal gyrus.
Figure 3: Regions showing significantly increased GM volume in non-responder (NR) patients compared to responder (R) patients. P: posterior, A: anterior, L: left, R: right. C: Cerebellum, PCC: Posterior Cingulate Cortex.