



# Feasibility of studying brain morphology in major depressive disorder with structural magnetic resonance imaging and clinical data from the electronic medical record: A pilot study

## Citation

Hoogenboom, Wouter S., Roy H. Perlis, Jordan W. Smoller, Qing Zeng-Treitler, Vivian S. Gainer, Shawn N. Murphy, Susanne E. Churchill, Isaac S. Kohane, Martha E. Shenton, and Dan V. Iosifescu. 2013. Feasibility of Studying Brain Morphology in Major Depressive Disorder with Structural Magnetic Resonance Imaging and Clinical Data from the Electronic Medical Record: A Pilot Study. *Psychiatry Research: Neuroimaging* 211, no. 3: 202–213. doi:10.1016/j.psychresns.2012.07.007.

## Published Version

doi:10.1016/j.psychresns.2012.07.007

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:28548727>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Published in final edited form as:

*Psychiatry Res.* 2013 March 30; 211(3): 202–213. doi:10.1016/j.psychresns.2012.07.007.

## Feasibility of studying brain morphology in major depressive disorder with structural magnetic resonance imaging and clinical data from the electronic medical record: A pilot study

Wouter S. Hoogenboom<sup>a,b,\*</sup>, Roy H. Perlis<sup>a,c</sup>, Jordan W. Smoller<sup>d</sup>, Qing Zeng-Treitler<sup>e,f</sup>, Vivian S. Gainer<sup>g</sup>, Shawn N. Murphy<sup>g,h</sup>, Susanne E. Churchill<sup>i</sup>, Isaac S. Kohane<sup>i</sup>, Martha E. Shenton<sup>b,j</sup>, and Dan V. Iosifescu<sup>a,k</sup>

<sup>a</sup>Depression Clinical and Research Program, Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, United States

<sup>b</sup>Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, United States

<sup>c</sup>Center for Human Genetic Research, Laboratory of Psychiatric Pharmacogenomics, Massachusetts General Hospital, Boston, MA 02114, United States

<sup>d</sup>Psychiatric Genetics Program in Mood and Anxiety Disorders, Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, United States

<sup>e</sup>University of Utah, Department of Biomedical Informatics, Salt Lake City, UT 84112, United States

<sup>f</sup>VA Salt Lake City Health Care System, Salt Lake City, UT, 84148, United States

<sup>g</sup>Information Systems, Partners HealthCare System, Inc., Charlestown, MA 02129, United States

<sup>h</sup>Laboratory of Computer Science, Massachusetts General Hospital, Boston, MA 02114, United States

<sup>i</sup>2b2 National Center for Biomedical Computing, Brigham and Women's Hospital, Boston, MA 02115, United States

<sup>j</sup>Clinical Neuroscience Division, Laboratory of Neuroscience, Department of Psychiatry, VA Boston Healthcare System, Brockton Division, Brockton, MA 02301 and Harvard Medical School, Boston, MA 02115, United States

<sup>k</sup>Mood and Anxiety Disorders Program, Mount Sinai School of Medicine, New York, NY 10029, United States

### Abstract

For certain research questions related to long-term outcomes or to rare disorders, designing prospective studies is impractical or prohibitively expensive. Such studies could instead utilize clinical and magnetic resonance imaging data (MRI) collected as part of routine clinical care, stored in the electronic medical record (EMR). Using major depressive disorder (MDD) as a

© 2012 Elsevier Ireland Ltd. All rights reserved.

\*Corresponding author: Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, 1249 Boylston Street, Boston, MA 02215, United States, Tel: +1 617 455 8929, Fax: +1 617 525 6150, hoogenboomw@gmail.com.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

disease model, we examined the feasibility of studying brain morphology and associations with remission using clinical and MRI data exclusively drawn from the EMR. Advanced automated tools were used to select MDD patients and controls from the EMR who had brain MRI data, but no diagnosed brain pathology. MDD patients were further assessed for remission status by review of clinical charts. Twenty MDD patients (eight full-remitters, six partial-remitters, and six non-remitters), and fifteen healthy control subjects met all study criteria for advanced morphometric analyses. Compared to controls, MDD patients had significantly smaller right rostral-anterior cingulate volume, and level of non-remission was associated with smaller left hippocampus and left rostral-middle frontal gyrus volume. The use of EMR data for psychiatric research may provide a timely and cost-effective approach with the potential to generate large study samples reflective of the real population with the illness studied.

## Keywords

MDD; refractory depression; structural MRI; hippocampus; legacy data; natural language processing

## 1. Introduction

While magnetic resonance imaging (MRI) studies with prospectively-acquired data make invaluable contributions to our understanding of underlying neuropathology in psychiatric disorders, such studies remain expensive and time-consuming. For research involving long-term follow-up or rare disorders prospective studies are very difficult or impractical. Using data available in the electronic medical record (EMR) may be a cost-effective solution for answering such research questions. This approach has advantages in efficiency and feasibility because the data has already been collected and the sheer size of electronic medical record systems may potentially result in large study samples. Moreover, study results may have increased ecological validity and reduced sampling bias.

Research in major depressive disorder (MDD) is one of the driving biology projects in which i2b2 (a national center for biomedical computing: <http://www.i2b2.org>) provides technical expertise (Murphy et al., 2007) to identify, process, and mine the EMR. Although the use and sharing of health information is associated with concerns related to patient privacy (Kohane and Altman, 2005), recent advances in the development of clinical data warehouses (Nalichowski et al., 2006), query tools (Holzbach et al., 2004; McDonald et al., 2005; Weber et al., 2009), and de-identification of patient information (Thomas et al., 2002; Gupta et al., 2004; Beckwith et al., 2006; Szarvas et al., 2007; Uzuner et al., 2007) have made clinical information more accessible for research while rigorously protecting patient privacy.

The use of EMR data is similar to the approach of combining legacy data from multiple research sites. For example, in a previous study (Fennema-Notestine et al., 2007) the feasibility of reanalyzing multi-site structural neuroimaging data was determined in elderly healthy subjects, and declining hippocampus volume across the age range was successfully replicated. The variability in data acquisition methods and subject samples across sites was minimized with careful data curation, image analysis, and statistical modeling.

In this study, our primary aim is to test the feasibility of using EMR data for neuropsychiatric research and to describe the specific methods required to detect research-quality clinical and brain MRI data in the EMR. Our secondary aim is to use this data to replicate morphological findings demonstrated previously in patients with major depression. In morphometric analyses we focus on the hippocampus, the anterior cingulate cortex

(ACC), and the rostral-middle frontal gyrus (rmfg). These structures are among the most studied and frequently reported brain regions to exhibit structural abnormalities in major depressive disorder (e.g., see reviews in Iosifescu and Lapidus, 2011; (Konarski et al., 2008; Koolschijn et al., 2009; Lorenzetti et al., 2009). We also evaluated long-term treatment outcome in MDD subjects, as a growing body of literature suggests that specific brain microstructural abnormalities (in both gray and white matter) underlie poor response to antidepressant treatments (Shah et al., 1998; Shah et al., 2002; Iosifescu et al., 2006; Sheline et al., 2010; Zhou et al., 2011).

In morphometric analyses, we hypothesized that hippocampus, ACC and rmfg gray matter volumes are lower in MDD subjects with poor long-term outcomes compared to remitted MDD subjects and to healthy control subjects. To our knowledge, the study reported here is the first to describe the systematic use of EMR data to study brain morphometry in major depressive disorder patients in association with long-term treatment outcome.

## 2. Methods

### 2.1 Subjects

For this retrospective chart review study, subjects were aged 16–75 years and selected from the electronic medical record system of Massachusetts General Hospital (MGH) or Brigham and Women’s Hospital (BWH) (see section 2.2 Data source) with the aid of computerized tools and diagnostic billing codes (see section 2.3 Natural Language Processing), a method previously described and validated by members of our research team (Perlis et al., 2012). All subjects in this study (except healthy controls) were previously diagnosed with major depressive disorder (MDD). To ensure the MDD diagnosis, we used the following selection methods: (1) billing data –that is, patients had at least one diagnosis of MDD (ICD-9 296.2x or 296.3x) recorded in the billing data or out-patient medical record at MGH or BWH, (2) natural language processing –that is, patients had at least one anti-depressant prescription and three or more outpatient psychiatry visit notes as determined by natural language processing (an automated method of chart review, see also 2.3 Natural language processing), (3) manual chart review –that is, electronic clinical charts were reviewed by an experienced board-certified clinical psychiatrist (DVI) to ensure MDD diagnosis.

Patients with billing codes for bipolar disorder (ICD-9 296.1x, 296.4x-296.9x), dementia/delirium (ICD-9 290.x), or psychotic disorder (ICD-9 295.x, 298.x) were excluded, as were those with other depressive disorders, such as dysthymia (ICD-9 300.4). Patients were also excluded if they had a past history of substance abuse/dependence or any diagnosis of neurological disorder, including brain tumors, stroke, dementia, status epilepticus, seizure disorder, CNS vasculitis and multiple sclerosis. Control subjects were free of any psychiatric (ICD-9 290–319, and Diagnosis Related Groups, DRG 424–432) or neurological disorder (ICD-9 320–389 and DRG 001–035, 524, 528–534, 543, 559, 737–739, 761–769).

### 2.2 Data source

Data in this study originated from an i2b2 “datamart”: A project-specific defined extract from the Research Patient Data Registry (RPDR; <http://rc.partners.org/rpdr>) by the i2b2 Workbench software (i2b2 v1.4, Boston, MA, USA) (Murphy et al., 2007). The RPDR is a centralized data warehouse that stores clinical information such as billing codes, sociodemographic data, and laboratory results acquired from electronic medical records of inpatient clinical systems, billing systems, laboratory systems, and outpatient electronic medical record systems at Partners HealthCare (Boston, MA, USA). Data in the RPDR currently exists for more than 4 million patients (Murphy et al., 2010). All aspects of this study were approved by the Institutional Review Board of Partners Healthcare System.

### 2.3 Natural Language Processing

Electronic medical records contain a wide range of patient information, in particular quantitative data such as dates of hospitalization, diagnostic billing codes, and patient demographics. However, qualitative information such as illness course or treatment outcome is not systematically entered in a structured format but often embedded as free-form typed text in physician notes. Natural Language Processing (NLP) is a technique for extracting such qualitative data by using researcher-defined criteria in the form of keywords and grammatical rules (Goryachev et al., 2006; Zeng et al., 2006). We used NLP to gather qualitative information from the complete EMR (i.e., data in structured format as well as in free-form typed text), including radiology reports, to verify the billing diagnosis of MDD (Perlis et al., 2012), and to verify the lack of diagnosable pathology in brain MRIs (i.e., read as “normal” or “unremarkable” in the radiology notes).

### 2.4 Chart review

After MDD patients were selected from the EMR system, they were assigned a ‘remission-status’ (i.e., full-remission, partial-remission, or non-remission) by an experienced research psychiatrist (DVI) by review of at least three psychiatry visit notes, blinded to MRI results. In more detail, we first determined the patient’s MRI exam date. Starting from this date, we went back in time up to 12 months to find psychiatry visit notes describing the patient’s remission status. The determination of the clinical gold standard for remission status is described elsewhere (Perlis et al., 2012). In short, remission status was determined by review of clinician notes available in the EMR describing the patient’s current mood state. The patient was classified as either “full-remitter” –that is, patient does not feel depressed and is back to normal, defined as absence or virtual absence of depressive symptoms throughout the 12 months interval preceding the MRI; “partial-remitter” –that is, patient shows fluctuating response or an intermediate state of depression, defined as partial response not sufficient for remission throughout the 12 months interval preceding the MRI; or “non-remitter” –that is, patient feels persistently depressed, defined as likely to meet criteria for a major depressive episode at any point during the 12 months interval preceding the MRI. The notes used could be from any point during the 12 months period. Thus, the two events (MRI exam and remission status) were not clearly related in time, and describe an associative rather than predictive relationship.

### 2.5 Image acquisition

MRI data for this study were collected from Massachusetts General Hospital (Boston, MA) as part of routine clinical treatment (Current Procedural Terminology codes, CPT 70336-70553) from 1999–2009 (to minimize variability in MR technology). We defined the following minimum acceptable MRI acquisition criteria: (1) T1-weighted data at 1.5 Tesla field strength, (2) slice thickness 1.6mm, (3) coronal acquisition, (4) complete field-of-view (i.e., full brain captured), and (5) good scan quality (i.e., high gray and white matter contrast and low motion artifacts).

### 2.6 Image processing

MRI data were transferred to and evaluated on a workstation at the Psychiatry Neuroimaging Laboratory (<http://pnl.bwh.harvard.edu/>) with the open-source software package 3DSlicer (<http://www.slicer.org>, Surgical Planning Laboratory, BWH, Boston, MA). Subjects with no T1 data or those with T1 data but not meeting the minimum image acquisition requirements were excluded from the study. T1 data were realigned using the line between the anterior and posterior commissures and the sagittal sulcus to correct for head tilt, and images were resampled into isotropic voxels of 0.9375 mm<sup>3</sup>. The FMRIB Software Library (FSL v.4.1; <http://www.fmrib.ox.ac.uk/>) was used for its brain extraction

tool (Smith, 2002) to assist in the removal of non-brain tissue from T1 data (skull stripping), and to estimate the intra-cranial volume to assist in controlling for individual differences in head size. Further, skull-stripped T1 data were processed with Freesurfer (v4.5.0; <http://surfer.nmr.mgh.harvard.edu/>) to produce automated volumetric measures of the hippocampus, anterior cingulate (rostral and dorsal sub regions) and rostral-middle frontal gyrus (see figure 1 for regions-of-interest). Freesurfer's definition of anatomical boundaries and methods used for the segmentation of (sub)cortical regions are previously described (Dale et al., 1999; Fischl et al., 2002; Desikan et al., 2006).

## 2.7 Automated segmentation versus manual tracing

Automated segmentation methods, including Freesurfer, have demonstrated high accuracy in volume measures compared to manual tracing (Fischl et al., 2002; Fischl et al., 2004b; Bergouignan et al., 2009). To provide additional validation of Freesurfer's automated method specific to our data, we manually traced the hippocampus as a proof of concept on a random subset of cases ( $n=10$ ), and we calculated the intra-correlation coefficient (ICC) to measure hippocampal volume agreement between manual and automated methods. Manual tracing of the hippocampus was performed by a trained investigator (WSH), blind to subjects' diagnostic status, and according to a procedure previously published by our laboratory (Shenton et al., 1992).

## 2.8 Statistical analysis

Group differences in demographic and clinical variables involving continuous data (age, intracranial volume) were analyzed with the Kruskal-Wallis test (MDD by remission status vs. controls), or Mann-Whitney U test (all MDD subjects vs. Controls). Non-parametric test procedures were used because of the small MDD groups and the non-normal distribution of the dependent variables. Differences in MRI exam indications, scan parameters, genders, and races between MDD patients and healthy controls were analyzed using  $\chi^2$ -tests.

Regions-of-interest (ROI) were corrected for differences in individual head-size, i.e., ROI as percentage of intra-cranial volume (ICV) = [absolute ROI volume/ICV] x100%. First, to test for volume differences between all MDD and control subjects we used independent-samples t-tests, and paired-samples t-tests to test within-subject cortical asymmetry. The magnitude of the difference was expressed by the effect size, Cohen's  $d$ . Second, Spearman's rho was used to quantify the linear relationship between gray matter volume (interval scale) and remission status (ordinal scale, i.e., healthy control, MDD full-remitter, MDD partial-remitter, and MDD non-remitter). All tests were conducted using a two-sided  $\alpha$ -level of 0.05, except for Spearman's rho where we used one-sided significance to test the linear relationship between volume and remission status given our pre-specified one-sided hypothesis: Lower gray matter volume was expected in association with non-remission, i.e. whether the MDD patient is fully remitted, partially remitted, or non-remitted.

The i2b2 datamart was used to collect information on antidepressant medications prescribed within 12 months prior to MRI acquisition (see supplemental Table 1). However, the variety of antidepressants prescribed and the small number of patients with available medication data precluded use of antidepressant treatment as a covariate in analyses. Results were considered significant at  $p<0.05$ , but no correction for multiplicity was used due to the small sample size and exploratory nature of this preliminary investigation. All analyses were conducted with SPSS 17.0 (Chicago, IL).

### 3. Results

A detailed summary of subject enrollment is presented in Figure 2. Subjects were step-wise excluded from the i2b2 datamart by application of inclusion and exclusion criteria until all study requirements were met.

#### 3.1 RPDR and NLP results: Major depressive disorder patients

The initial RPDR query to create the MDD data mart was performed in December of 2007, and resulted in 123369 patients with at least a single episode of major depressive disorder, of which 5224 had concurrent MRI data and no history of comorbid mental or neurologic disorder. Natural language processing subsequently selected 880 patients with brain MRI data explicitly described as “normal” or “unremarkable” in the radiology notes, including 320 patients with multiple clinical notes available describing the patient’s status.

#### 3.2 RPDR and NLP results: Healthy control subjects

Separate RPDR queries were performed in February and May of 2009, and resulted in a datamart comprising 14791 individuals with MRI data but with a history free of any mental and neurologic disorder. This pool was defined as the ‘healthy control data mart’ and used to identify controls from who matched well to the study-eligible MDD patients. The concept for the matching algorithm has been described in detail elsewhere (Murphy et al., 2010). In short, matching was based on patient demographics (i.e., gender, age within 5 years, and race) and the number of ‘observational facts’. Observational facts refer to any coded attribute about the patient available in the EMR, such as a code for disease (e.g., diabetes), a code for a specific medication the patient is on (e.g., insulin), or a code for an ordered lab test (e.g., cholesterol). In general, the sicker a patient is the more observational facts the patient has; therefore, the number of facts can be used as a proxy for healthcare utilization (and to ensure a comparable level of concomitant disease history between subjects that is not related to depression). Control subjects were only included if an individual match was identified based on patient demographics, and observational facts within 50%, i.e., controls had no more or less than twice the amount of healthcare utilization than MDD patients. From the healthy control data mart we identified 869 matches to the study-eligible MDD patients. These control matches were further analyzed for the presence of brain pathology with natural language processing: 159 were found with normal brain MRI data.

#### 3.3 MRI data eligibility

320 MDD patients and 159 matched control subjects with normal brain MRI data were further evaluated for the presence of high resolution 1.5T coronal T1-weighted MRI acquisitions.

We found fifty-three MDD patients (17%) and twenty matched control subjects (13%) with high resolution MRI acquisitions (i.e., slice thickness  $\leq 1.6$ mm). Six subjects had medium resolution (i.e., slice thickness  $>1.6$ mm but  $<5$ mm, 1%), and all other subjects ( $n=400$ , 84%) had low image resolution (i.e., slice thickness  $> 5$ mm).

A significant, positive correlation (Spearman’s  $r=0.87$ ,  $p=0.003$ ) was found between the number of subjects with high resolution T1 data and the year of acquisition, which suggests there was a growing trend of high resolution MRI use in clinical practice in recent years.

However, only half of all high resolution scans ( $n=35$ , 48%) also met further criteria for image quality, clinical, and demographic eligibility. Of those excluded, twelve had magnetic field strengths not equal to 1.5T, ten had the brain not fully captured (i.e., 2 cases only had edges of brain lobes missing, whereas in 8 cases only the midbrain was visible), four had

low quality MRI (i.e., low image contrast or motion artifacts), six had severe comorbidities, and six were underage (below 16 years) (see also bottom of flowchart, Figure 2). MRI examples encountered in this study are presented in Figure 3.

### 3.4 Chart review of clinical status

At this point, 20 datasets from MDD subjects (aged  $42.7 \pm 17.6$  years; 9 females) and 15 from healthy subjects (aged  $32.1 \pm 13.5$  years; 4 females) were identified that met all study criteria. By review of clinical charts, MDD patients were classified as follows: 8 were full-remitters (aged  $39.0 \pm 12.9$  years; 4 females), 6 were partial-remitters (aged  $44.9 \pm 17.1$  years; 2 females), and 6 were non-remitters (aged  $45.6 \pm 24.8$  years; 3 females). Patients and healthy controls did not differ significantly in age, gender, race, reason for MRI, or intra-cranial volume (Table 1).

**3.4.1 Reasons for MRI exam**—Subjects in this study underwent MRI examination for several reasons. The majority had MRIs to rule out brain pathology after episodes of syncope, fainting, or dizziness ( $n=18$ , 51%). Others with a mix of signs and symptoms not specific to a particular diagnosis received “routine imaging” ( $n=8$ , 23%) to rule out any form of brain abnormality. Some patients had specific clinical symptoms, such as transient aphasia, ataxia, vertigo, altered mental state, one-side sensation loss or numbness of body parts that required MRI follow-up to rule out possible stroke, infarction, or aneurysm ( $n=5$ , 14%). Others had primary symptoms of chronic headaches and/or migraines ( $n=3$ , 9%), whereas one patient ( $n=1$ , 3%) was examined to rule out CNS metastases of colon cancer. Each subject’s exam was reviewed by a clinical radiologist at the time of scanning and concluded as “normal” or “unremarkable MRI of the brain” in the radiology report. MRI exam indications did not differ between MDD subjects vs. controls ( $\chi^2=3.21$ ,  $df=4$ ,  $P=0.52$ ), or between MDD groups classified by remission status vs. controls ( $\chi^2=14.08$ ,  $df=12$ ,  $P=0.30$ ).

### 3.5 Pulse sequences

A detailed description of pulse sequences is provided in Table 2. The majority of subjects had images that were 3-dimensional spoiled-gradient recalled sequence (SPGR; repetition time, 30 milliseconds; echo time, 4.85 milliseconds; number of acquisitions, 1; field of view, 22cm; matrix, 256×256 pixels; yielding 124 contiguous coronal sections with a defined voxel size of 0.86×0.86×1.5 mm). Image acquisition parameters were not significantly different between study groups (Table 2).

### 3.6 Morphometric analyses

Subcortical and surface-based cortical segmentations were visually inspected for accuracy prior to inclusion in this study, and only minimal corrections were made to remove or replace tissue.

**3.6.1 Reliability analyses**—Excellent agreement between Freesurfer’s automated neuroanatomical labeling and manual tracing was observed for left ( $ICC=0.87$ ,  $P<0.01$ ) and right hippocampus ( $ICC=0.95$ ,  $P<0.0001$ ). This result is in accordance with a previous test of agreement between the volumetric methods in chronic major depressive disorder (Tae et al., 2008), and suggests that the automated method, which is time-efficient and rater-independent, can be reliably used for clinical MRI data such as collected in our study.

**3.6.2 Major depressive disorder patients**—Morphometric results are presented in Table 3 and Figure 4. Compared to the control group, MDD subjects had bilateral lower absolute and relative (head-size corrected) volumes for the hippocampus, rostral-ACC, and



rostral-middle frontal gyrus, but not caudal-ACC, statistically significant for the right rostral-ACC ( $t=2.07$ ,  $df=33$ ,  $P=0.046$ , Cohen's  $d=0.53$ ).

Paired-samples t-tests also revealed significant left dominant hippocampus asymmetry for healthy control subjects (left greater than right;  $P=0.02$ , Cohen's  $d=0.23$ ), but not for MDD patients ( $P=0.97$ ). Both groups had left dominant asymmetry for the rostral-ACC (control subjects:  $P=0.03$ , Cohen's  $d=0.97$ ; MDD patients:  $P<0.001$ , Cohen's  $d=1.44$ ), whereas caudal-ACC and rostral-middle frontal gyrus revealed right greater than left volumes for MDD patients (caudal-ACC:  $P=0.03$ , Cohen's  $d=0.67$ ; rostral-middle frontal gyrus:  $P=0.03$ , Cohen's  $d=0.45$ ) but not for control subjects (caudal-ACC:  $P=0.13$ ; rostral-middle frontal gyrus:  $P=0.09$ ).

**3.6.3 Major depressive disorder patients classified by remission status—**We did not detect differences in regions-of-interest among MDD patients classified by remission status, and controls (Kruskall-Wallis H values  $<5.28$ ,  $df=3$ ,  $P>0.15$ ), most likely due to the low statistical power as a result of small group sizes and high variability. However, non-parametric Spearman correlations suggested that non-remission was significantly associated with head-size corrected smaller volumes of the left hippocampus ( $r_{35}=-0.32$ ,  $P=0.03$ ), and left rostral-middle frontal gyrus ( $r_{35}=-0.35$ ,  $P=0.02$ ), whereas statistical significance was reached for the right hippocampus ( $r_{35}=-0.28$ ,  $P=0.053$ ), right rostral-middle frontal gyrus ( $r_{35}=-0.26$ ,  $P=0.07$ ), and right rostral-ACC ( $r_{35}=-0.26$ ,  $P=0.07$ ), but not caudal-ACC (left:  $r_{35}=0.02$ ,  $P=0.46$ ; right:  $r_{35}=0.08$ ,  $P=0.33$ ) and left rostral-ACC ( $r_{35}=0.01$ ,  $P=0.48$ ) (Table 4 and Figure 5).

## 4. Discussion

In this pilot study, as a principal aim, we demonstrate the feasibility as well as the challenges of using the EMR system in identifying MDD patients with concurrent brain MRI data to study questions related to brain morphology and its association to remission status in major depression. Although the initial yield of MRI data was relatively low in this novel approach, this work is important from methodological perspective as it shows that a study like this can be done. Moreover, as data from EMR systems continue to grow in numbers and availability, future studies using legacy data are expected to yield larger sample sizes. As a secondary aim, we demonstrate subtle group differences and trends in the data, despite variability in data acquisition and relatively small sample sizes used for morphometric analysis. The morphometric results are in accordance with previous findings, including an 8–10% volume reduction of bilateral hippocampus (Videbeck and Ravnkilde, 2004), and reduced anterior cingulate volume (Caetano et al., 2006) in patients with unipolar depression, compared to non-depressed individuals. Also, the presence or absence of brain structure asymmetry may give us insight into the neurobiology of mental disorders (Wang et al., 2001; Pedraza et al., 2004). Finally, the volume of the hippocampus was associated with non-remission (i.e., those who were partially or non-remitted were more likely to have smaller volumes compared to normal controls or fully-remitted MDD patients), which is in accordance with previous findings suggesting that microstructural abnormalities of the hippocampus may be indicative for increased vulnerability to treatment-resistance (Vakili et al., 2000; Zhou et al., 2011). Given the feasibility and limitations of our data, the morphometric results should be interpreted cautiously and are secondary to the methods providing indirect proof of concept.

### 4.1 Potential benefits and pitfalls of our approach

Using clinical and MRI data originally collected as part of routine clinical care represents an alternative approach to traditional data collection in neuroimaging research. This approach has advantages in efficiency and feasibility because the data has already been collected, the

sheer size of electronic medical record systems may potentially result in large study samples, and the results may be achieved in a more expeditive and cost-effective manner. Calculations show that reusing clinical data can result in significant cost reduction in regard to sample acquisition and phenotyping in genomic research (Murphy et al., 2009). Similarly, reusing MRI data collected as part of clinical care may have cost effective consequences for neuroimaging research considering the high costs of MRI acquisition. Also, the ability to reliably ascertain patients' diagnoses based on the full EMR (including free text notes) has been validated for rheumatoid arthritis (Liao et al., 2010) and major depressive disorder (Perlis et al., 2012).

On the other hand, there are several pitfalls related to our approach. First, we relied on clinical characterization based on review of medical records, where absent or insufficient data in the EMR is prone to some level of misclassification of a patient's remission status.

Second, subjects in our sample may not represent a random sample of the population, simply because they received an MRI (even though the reason for MRI was random, i.e., not related to the depression and not biased by group). For example, individuals with higher socioeconomic status (SES) may be more likely to utilize imaging modalities such as MRI (Demeter et al., 2005; You et al., 2009); therefore, our sample may be underrepresented for people from lower socioeconomic class. However, whether or not SES is confounded with the probability of receiving an MRI is deferred to future work, because information on SES was not systematically collected from the EMR. Further, people who received an MRI represent only a small proportion of the population. Even though we only used MRIs that were assessed as normal or unremarkable, as described in the radiology report by the clinical radiologist, this does not assure complete lack of pathology. Thus, it is possible that our sample represents the most clinically symptomatic patients. These types of subject selection bias are inherent to a study with legacy MRI data and represent a limitation to our approach. As a growing number of healthcare institution adopt EMR systems, future data-mining studies may profit from a broader range of available patient information, enabling the identification of more representative population samples. Of note, the fact that our approach involves "real patients" and not just volunteers in a research study increases the ecological validity of our study, and is one of the strengths of our approach.

Third, the absence of a clearly defined temporal relationship between the time of MRI and the time of the patient's remission status assessment was unavoidable given that the collection of legacy data was part of ongoing and routine clinical care. However, we do not believe that this influenced the morphometric results to a great extent given that major changes in brain structure are unlikely to appear in short time intervals, i.e., in a few months. Nonetheless, the applicability of our approach to study other brain regions and other psychiatric or medical populations should be assessed.

## 4.2 Study limitations and future perspective

The main limitation to our study finding is the relatively low yield of MRI data given the size of the parent database. Therefore, this pilot study should be considered to be more of a small-scale proof of concept, as the novel approach presented here requires further development in the efficiency of the data-mining process. However, we note that some numbers found in our study are specific to our study research question and will not necessarily be the same for other studies. For example, if we were only interested in finding MDD patients and not further investigating remission status, then we would use the pool comprising 880 MDD patients rather than 320 (see figure 2, patient enrolment flowchart). In that case, we would nearly triple the subject pool and likely increase our final sample size. Furthermore, although the parent database in our study was relatively large with over 4 million patients, it has been suggested that a database of 40–50 million patients is ideal

(Tannen et al., 2009). There are several reasons to expect that EMR datasets will continue to grow. First, a greater number of institutions have adopted EMR systems in recent years (Jha et al., 2009). Second, a greater number of institutions are making EMR data available for research (Murphy et al., 2010). Third, as clinical practices continue to upgrade their MRI systems, larger numbers derived from the EMR will meet the high quality-standards of research MRI data. For example, recent developments have led the move of scanner upgrades, such as in magnetic field strength from 1.5T to 3T, from a research setting to clinical practice (Tanenbaum, 2006). This development is in accordance to our finding of increased availability of high resolution MRI data across clinical practices in recent years. Even with a small percentage of clinical and MRI data eligible for research, it is the sheer size and the growing availability of EMR data as well as continued efforts and advancements in data-mining that will result in larger study samples. Of note, we already demonstrated the ability to identify larger study cohorts from the current database to perform quantitative analyses in a diffusion MRI study, yielding 92 MDD patients and 58 normal controls (Iosifescu et al., 2010; Hoogenboom et al., 2012). This suggests that the type of MRI data may also play a role—that is, in the EMR system, diffusion weighted MRI may be more abundant or closer to research standards than anatomical MRI.

Another study limitation is that MRI data has been previously collected as part of routine clinical treatment using multiple scanner hardware and imaging protocols, which may affect the reliability of measures and reproducibility of results (Kruggel et al., 2010). Although it has been demonstrated that volumes derived from automated segmentation of structural data can be reliably measured within the same scanner platform, even after upgrades, combining data across sites remains problematic (Han et al., 2006; Jovicich et al., 2009). We limited the variability in image acquisitions by selecting T1 data with similar acquisition parameters and those with high variability were excluded. We note that variability in image acquisition was primarily within, and not between study groups. Therefore, it was expected and found in this study that the variability in image acquisitions was randomly distributed and similar among diagnostic groups. One concern, however, with such increased variability is that it increases variability in the outcome measures, and thus lowers the study statistical power.

Lastly, controlling for medication status is very complicated in a data-mining EMR study because the data is collected in advance as part of ongoing clinical care in which the type, dosage, and administration of medications is unique to each patient. Also, the lack of statistical power did not allow us to use medication status as a covariate, or to restrict the number of different medications by excluding more patients given our small sample size. This potential confound represents a limitation to our approach, as well as to our study results although we have found no effects of chronic selective serotonin reuptake inhibitor (SSRI) exposure on brain structure, in a larger data-mined diffusion MRI study (Hoogenboom et al., 2012). Our study results should be seen as preliminary and future studies with legacy data may be able to focus on MRI explorations before or at the very early phases of treatment to avoid confounding effects of medication and to derive brain image predictors of therapeutic response.

## 5. Conclusion

In this pilot study we used the EMR to acquire research-quality clinical and brain MRI data from MDD patients to study brain morphology and its association to remission status. Our results demonstrate the feasibility as well as the challenges related to this novel approach to data collection. Morphometric findings are replicated that were previously demonstrated in patients with major depression, including the hippocampus' vulnerability to treatment-resistance. Despite the limitations of using EMR data, with continuing efforts to overcome computational challenges intrinsic to medical record systems, and initiatives like Harvard

Catalyst ([www.catalyst.harvard.edu](http://www.catalyst.harvard.edu)) to acquaint clinicians and scientists with tools and technologies in regard to translational imaging, plus improvements in brain segmentation algorithms to reduce site-specific variance (Fischl et al., 2004a), there is great potential to use the EMR as a rich resource of clinical and MRI data for translational neuroimaging research. This approach will not replace studies with prospectively acquired data, but rather complement the methods for data acquisition to study questions of scientific interest related to psychiatric populations, especially in situations where prospective studies may be impractical, time-consuming, or prohibitively expensive.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work is funded as a Driving Biological Project (PI: Dr. Iosifescu) by the i2b2 grant (National Institutes of Health, NIH U54-LM008748; PI: Dr. Kohane). Dr. Iosifescu is supported by the National Institutes of Health Career Development Award K23 MH067111. Dr. Shenton is supported by grants from the National Institutes of Health (R01 MH 50740), and the Department of Veterans Affairs (VA) (VA Merit Award and VA Schizophrenia Research Center Grant). Dr. Perlis is supported by a grant from the National Institute of Mental Health (MH086026). The authors thank Victor Castro, B.S., for his assistance in collecting data from the Research Patient Data Registry; Sergey Goryachev, M.S., for his assistance with natural language processing; Kenneth Nesbitt, for providing access to magnetic resonance imaging data from Massachusetts General Hospital; Dr. Bouix and Dr. Kubicki, for their advice in the development of the imaging study; and special thanks to all research assistants at the psychiatry neuroimaging laboratory for their kind help at the time of this study.

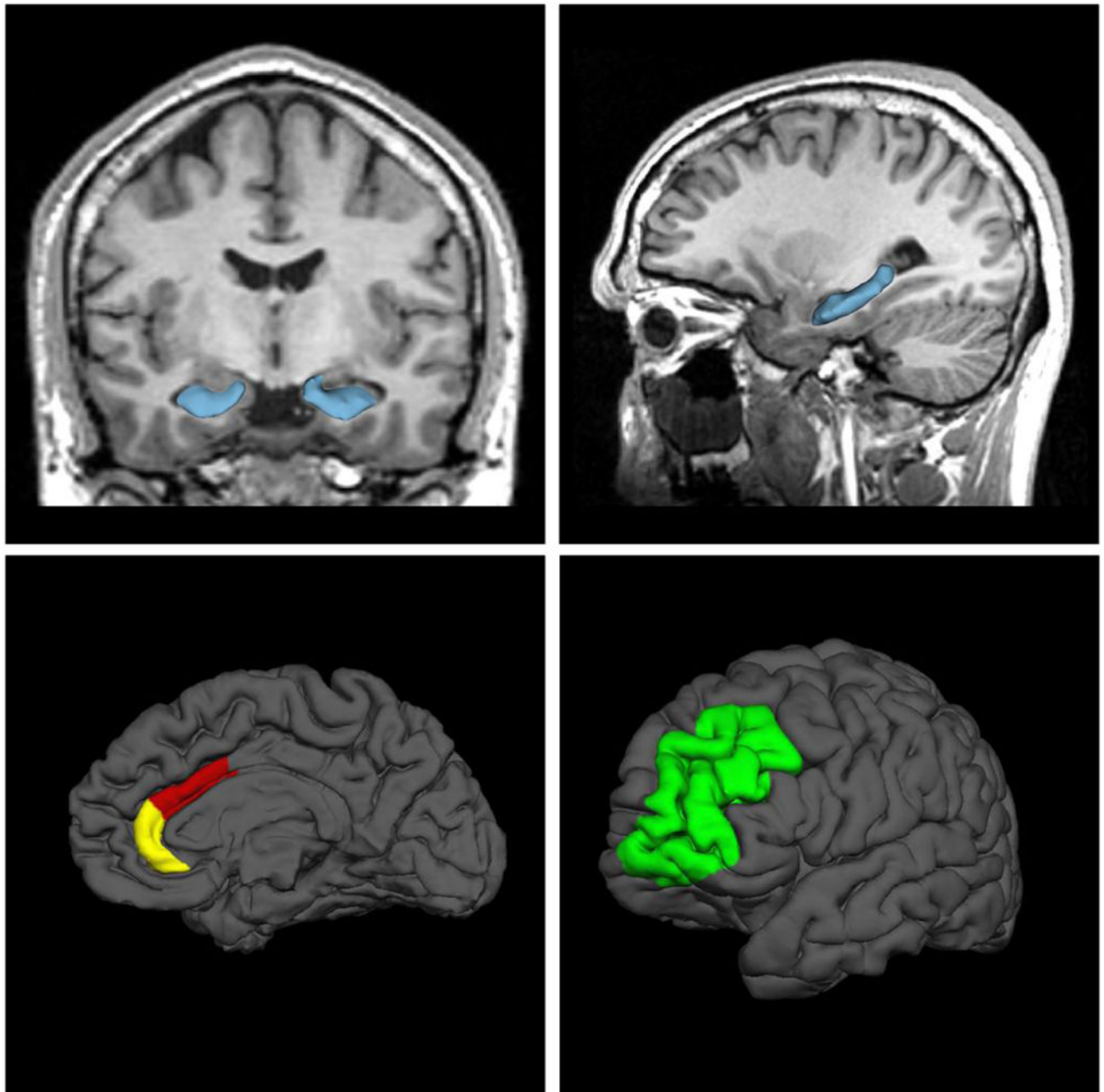
## References

- Beckwith BA, Mahaadevan R, Balis UJ, Kuo F. Development and evaluation of an open source software tool for deidentification of pathology reports. *Biomedcentral Medical Informatics and Decision Making*. 2006; 6:12.
- Bergouignan L, Chupin M, Czechowska Y, Kinkingnehun S, Lemogne C, Le Bastard G, Lepage M, Garnero L, Colliot O, Fossati P. Can voxel based morphometry, manual segmentation and automated segmentation equally detect hippocampal volume differences in acute depression? *Neuroimage*. 2009; 45(1):29–37. [PubMed: 19071222]
- Caetano SC, Kaur S, Brambilla P, Nicoletti M, Hatch JP, Sassi RB, Mallinger AG, Keshavan MS, Kupfer DJ, Frank E, Soares JC. Smaller cingulate volumes in unipolar depressed patients. *Biological Psychiatry*. 2006; 59(8):702–706. [PubMed: 16414029]
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*. 1999; 9(2):179–194. [PubMed: 9931268]
- Demeter S, Reed M, Lix L, MacWilliam L, Leslie WD. Socioeconomic status and the utilization of diagnostic imaging in an urban setting. *Canadian Medical Association journal*. 2005; 173(10):1173–1177. [PubMed: 16275968]
- Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006; 31(3): 968–980. [PubMed: 16530430]
- Fennema-Notestine C, Gamst AC, Quinn BT, Pacheco J, Jernigan TL, Thal L, Buckner R, Killiany R, Blacker D, Dale AM, Fischl B, Dickerson B, Gollub RL. Feasibility of multi-site clinical structural neuroimaging studies of aging using legacy data. *Neuroinformatics*. 2007; 5(4):235–245. [PubMed: 17999200]
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002; 33(3):341–355. [PubMed: 11832223]

- Fischl B, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT, Dale AM. Sequence-independent segmentation of magnetic resonance images. *Neuroimage*. 2004a; 23(Suppl 1):S69–84. [PubMed: 15501102]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. *Cerebral Cortex*. 2004b; 14(1):11–22. [PubMed: 14654453]
- Goryachev, S.; Sordo, M.; Zeng, QT. A suite of natural language processing tools developed for the I2B2 project. American Medical Informatics Association Annual Symposium Proceedings; 2006. p. 931
- Gupta D, Saul M, Gilbertson J. Evaluation of a deidentification (De-Id) software engine to share pathology reports and clinical documents for research. *American Journal of Clinical Pathology*. 2004; 121(2):176–186. [PubMed: 14983930]
- Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, Busa E, Pacheco J, Albert M, Killiany R, Maguire P, Rosas D, Makris N, Dale A, Dickerson B, Fischl B. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage*. 2006; 32(1):180–194. [PubMed: 16651008]
- Holzbach A, Porter A, Kohane I, Berkowicz D. A query engine for distributed medical databases. *Medinfo*. 2004; 21:1519.
- Hoogenboom WS, Perlis RH, Smoller JW, Zeng-Treitler Q, Gainer VS, Murphy SN, Churchill SE, Kohane IS, Shenton ME, Iosifescu DV. Limbic system white matter microstructure and long-term treatment outcome in major depressive disorder: A diffusion tensor imaging study using legacy data. *World Journal of Biological Psychiatry*. 2012;10.3109/15622975.2012.669499
- Iosifescu DV, Hoogenboom WS, Perlis RH, Smoller JW, Zeng-Treitler Q, Gainer VS, Murphy SN, Churchill SE, Kohane I, Shenton ME. Limbic system white matter microstructure and treatment outcome in major depressive disorder: A large scale DTI study using legacy data. *Neuropsychopharmacology*. 2010; 35:S78–187. A131.
- Iosifescu DV, Renshaw PF, Lyoo IK, Lee HK, Perlis RH, Papakostas GI, Nierenberg AA, Fava M. Brain white-matter hyperintensities and treatment outcome in major depressive disorder. *British Journal of Psychiatry*. 2006; 188:180–185. [PubMed: 16449707]
- Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR, Ferris TG, Shields A, Rosenbaum S, Blumenthal D. Use of electronic health records in U.S. hospitals. *New England Journal of Medicine*. 2009; 360(16):1628–1638. [PubMed: 19321858]
- Jovicich J, Czanner S, Han X, Salat D, van der Kouwe A, Quinn B, Pacheco J, Albert M, Killiany R, Blacker D, Maguire P, Rosas D, Makris N, Gollub R, Dale A, Dickerson BC, Fischl B. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage*. 2009; 46(1):177–192. [PubMed: 19233293]
- Kohane IS, Altman RB. Health-information altruists—a potentially critical resource. *New England Journal of Medicine*. 2005; 353(19):2074–2077. [PubMed: 16282184]
- Konarski JZ, McIntyre RS, Kennedy SH, Rafi-Tari S, Soczynska JK, Ketter TA. Volumetric neuroimaging investigations in mood disorders: bipolar disorder versus major depressive disorder. *Bipolar Disorder*. 2008; 10(1):1–37.
- Koolschijn PC, van Haren NE, Lensvelt-Mulders GJ, Hulshoff Pol HE, Kahn RS. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Human Brain Mapping*. 2009; 30(11):3719–3735. [PubMed: 19441021]
- Kruggel F, Turner J, Muftuler LT. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *Neuroimage*. 2010; 49(3):2123–2133. [PubMed: 19913626]
- Liao KP, Cai T, Gainer V, Goryachev S, Zeng-treitler Q, Raychaudhuri S, Szolovits P, Churchill S, Murphy S, Kohane I, Karlson EW, Plenge RM. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis Care & Research*. 2010; 62(8):1120–1127. [PubMed: 20235204]
- Lorenzetti V, Allen NB, Fornito A, Yucel M. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *Journal of Affective Disorders*. 2009; 117(1–2): 1–17. [PubMed: 19237202]

- McDonald, CJ.; Dexter, P.; Schadow, G.; Chueh, HC.; Abernathy, G.; Hook, J.; Blevins, L.; Overhage, JM.; Berman, JJ. SPIN query tools for de-identified research on a humongous database. American Medical Informatics Association Annual Symposium Proceedings; 2005. p. 515-519.
- Murphy S, Churchill S, Bry L, Chueh H, Weiss S, Lazarus R, Zeng Q, Dubey A, Gainer V, Mendis M, Glaser J, Kohane I. Instrumenting the health care enterprise for discovery research in the genomic era. *Genome Research*. 2009; 19(9):1675–1681. [PubMed: 19602638]
- Murphy, SN.; Mendis, M.; Hackett, K.; Kuttan, R.; Pan, W.; Phillips, LC.; Gainer, V.; Berkowicz, D.; Glaser, JP.; Kohane, I.; Chueh, HC. Architecture of the open-source clinical research chart from Informatics for Integrating Biology and the Bedside. American Medical Informatics Association Annual Symposium Proceedings; 2007. p. 548-552.
- Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, Kohane I. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *Journal of the American Medical Informatics Association*. 2010; 17(2):124–130. [PubMed: 20190053]
- Nalichowski, R.; Keogh, D.; Chueh, HC.; Murphy, SN. Calculating the benefits of a Research Patient Data Repository. American Medical Informatics Association Annual Symposium Proceedings; 2006. p. 1044
- Pedraza O, Bowers D, Gilmore R. Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. *Journal of the International Neuropsychological Society*. 2004; 10(5):664–678. [PubMed: 15327714]
- Perlis RH, Iosifescu DV, Castro VM, Murphy SN, Gainer VS, Minnier J, Cai T, Goryachev S, Zeng Q, Gallagher PJ, Fava M, Weilburg JB, Churchill SE, Kohane IS, Smoller JW. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. *Psychological Medicine*. 2012; 42(1):41–50. [PubMed: 21682950]
- Shah PJ, Ebmeier KP, Glabus MF, Goodwin GM. Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression. Controlled magnetic resonance imaging study. *British Journal of Psychiatry*. 1998; 172:527–532. [PubMed: 9828995]
- Shah PJ, Glabus MF, Goodwin GM, Ebmeier KP. Chronic, treatment-resistant depression and right fronto-striatal atrophy. *British Journal of Psychiatry*. 2002; 180:434–440. [PubMed: 11983641]
- Sheline YI, Pieper CF, Barch DM, Welsh-Bohmer K, McKinstry RC, MacFall JR, D'Angelo G, Garcia KS, Gersing K, Wilkins C, Taylor W, Steffens DC, Krishnan RR, Doraiswamy PM. Support for the vascular depression hypothesis in late-life depression: results of a 2-site, prospective, antidepressant treatment trial. *Archives of General Psychiatry*. 2010; 67(3):277–285. [PubMed: 20194828]
- Shenton ME, Kikinis R, Jolesz FA, Pollak SD, LeMay M, Wible CG, Hokama H, Martin J, Metcalf D, Coleman M, McCarley RW. Abnormalities of the left temporal lobe and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. *New England Journal of Medicine*. 1992; 327(9):604–612. [PubMed: 1640954]
- Smith SM. Fast robust automated brain extraction. *Human Brain Mapping*. 2002; 17(3):143–155. [PubMed: 12391568]
- Szarvas G, Farkas R, Busa-Fekete R. State-of-the-art anonymization of medical records using an iterative machine learning framework. *Journal of the American Medical Informatics Association*. 2007; 14(5):574–580. [PubMed: 17823086]
- Tae WS, Kim SS, Lee KU, Nam EC, Kim KW. Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder. *Neuroradiology*. 2008; 50(7):569–581. [PubMed: 18414838]
- Tanenbaum LN. Clinical 3T MR imaging: mastering the challenges. *Magnetic Resonance Imaging Clinics of North America*. 2006; 14(1):1–15. [PubMed: 16530631]
- Tannen RL, Weiner MG, Xie D. Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings. *British Medical Journal*. 2009; 338:b81. [PubMed: 19174434]
- Thomas, SM.; Mamlin, B.; Schadow, G.; McDonald, C. A successful technique for removing names in pathology reports using an augmented search and replace method. American Medical Informatics Association Annual Symposium Proceedings; 2002. p. 777-781.

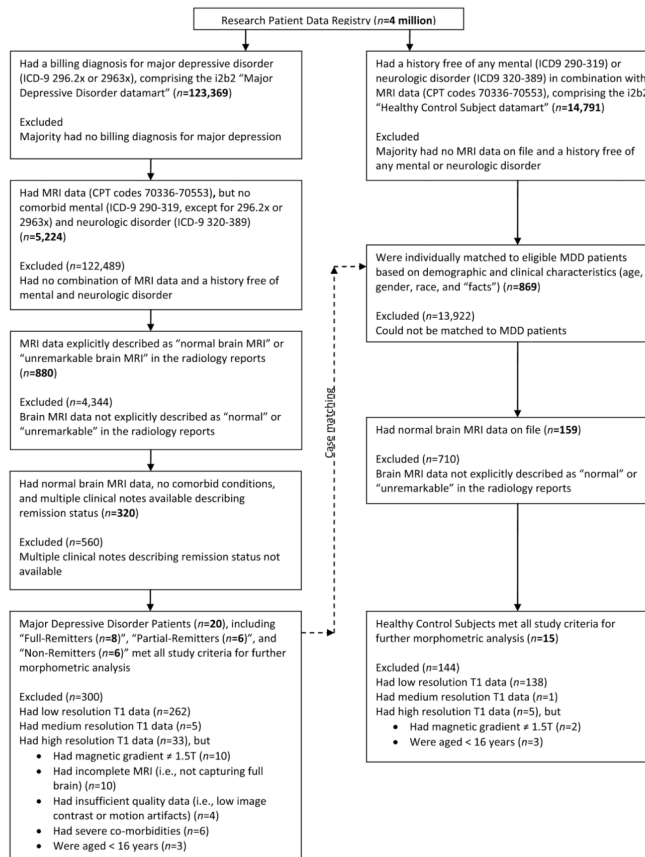
- Uzuner O, Luo Y, Szolovits P. Evaluating the state-of-the-art in automatic de-identification. *Journal of the American Medical Informatics Association*. 2007; 14(5):550–563. [PubMed: 17600094]
- Vakili K, Pillay SS, Lafer B, Fava M, Renshaw PF, Bonello-Cintron CM, Yurgelun-Todd DA. Hippocampal volume in primary unipolar major depression: a magnetic resonance imaging study. *Biological Psychiatry*. 2000; 47(12):1087–1090. [PubMed: 10862809]
- Videbech P, Ravnkilde B. Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry*. 2004; 161(11):1957–1966. [PubMed: 15514393]
- Wang L, Joshi SC, Miller MI, Csernansky JG. Statistical analysis of hippocampal asymmetry in schizophrenia. *Neuroimage*. 2001; 14(3):531–545. [PubMed: 11506528]
- Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, Kohane IS. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. *Journal of the American Medical Informatics Association*. 2009; 16(5):624–630. [PubMed: 19567788]
- You JJ, Venkatesh V, Laupacis A. Better access to outpatient magnetic resonance imaging in Ontario - But for whom? *Open Medicine*. 2009; 3(1):e22–25. [PubMed: 19946389]
- Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, comorbidity and smoking status for asthma research: evaluation of a natural language processing system. *Biomedcentral Medical Informatics and Decision Making*. 2006; 6:30.
- Zhou Y, Qin LD, Chen J, Qian LJ, Tao J, Fang YR, Xu JR. Brain microstructural abnormalities revealed by diffusion tensor images in patients with treatment-resistant depression compared with major depressive disorder before treatment. *European Journal of Radiology*. 2011; 80(2):450–454. [PubMed: 20634013]



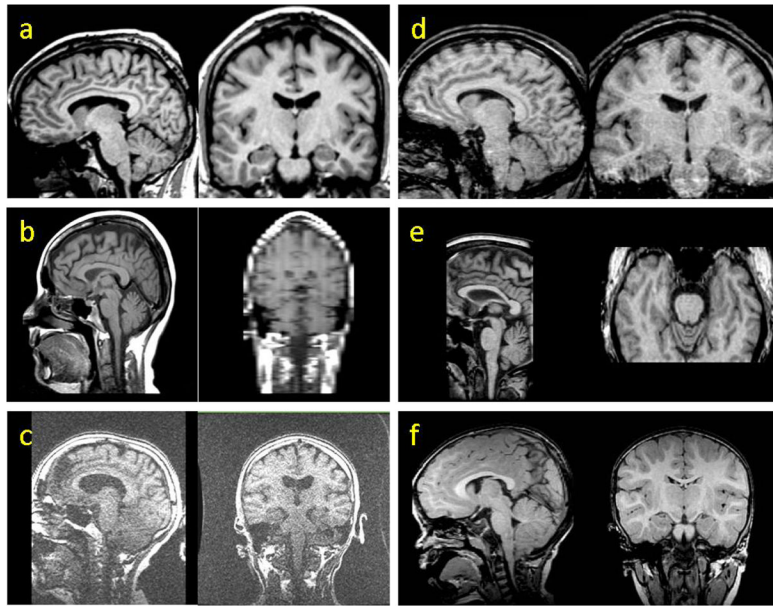
**Fig. 1.**

Regions-of-interest in this study: hippocampus (top row; left figure = anterior hippocampus in blue in frontal view of the brain; right figure = whole body of the hippocampus in blue in sagittal view of the brain; model reconstructions in 3D Slicer), anterior cingulate cortex (bottom left figure; ACC in medial view of Freesurfer reconstructed brain: lower portion in yellow represents rostral ACC and upper portion in red represents caudal ACC), and rostral-middle frontal gyrus (bottom right figure; rmfg in green in lateral view of Freesurfer reconstructed brain).

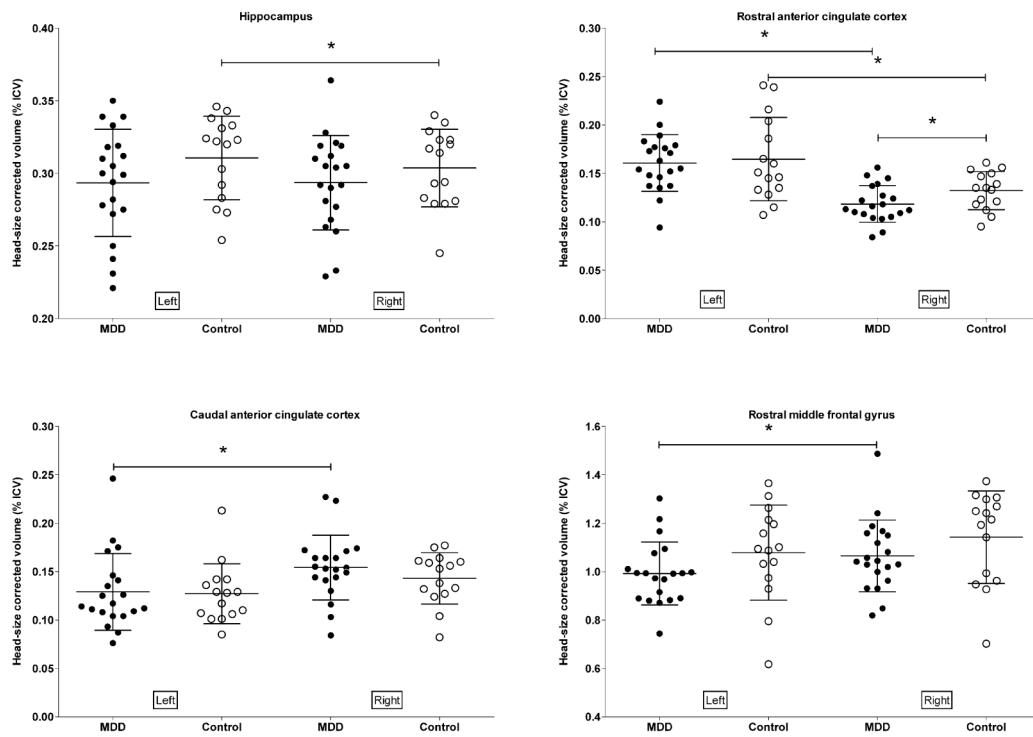




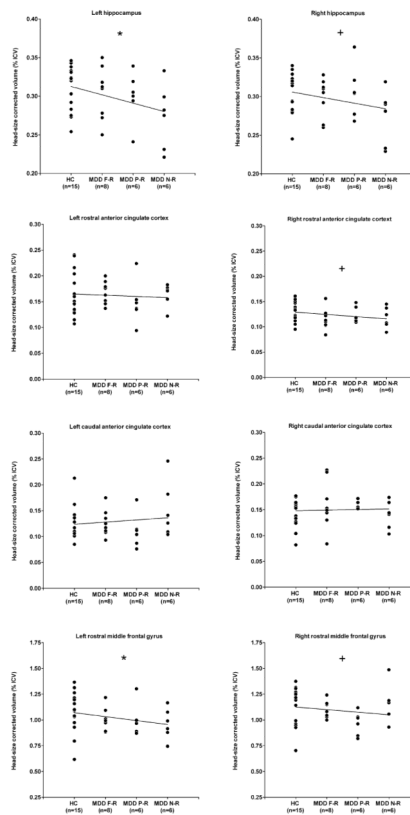
**Fig. 2.** Patient enrolment flowchart.



**Fig. 3.** Variability of image resolution, scan quality, and demographics of subjects with T1-weighted normal brain MRIs encountered in this study. Figure (a) represents a subject with high resolution, good quality scan who was enrolled in this study. Examples of subjects whose data were ineligible because of (b) low out-of-plane resolution (slice thickness=5mm), (c) low image contrast, (d) motion artifacts, (e) incomplete scan acquisition, (f) age = 16 years. Images represent sagittal (left) and coronal (right) plane, except for (e), which is sagittal (left) and axial (right) plane.



**Fig. 4.** Head-size corrected regions-of-interest for all MDD and matched control subjects. Each symbol represents one individual. Bars represent mean±SD. Compared to control subjects, MDD patients had head size-corrected bilateral lower volumes of the hippocampus, rostral-anterior cingulate, and rostral-middle frontal gyrus, but not caudal-anterior cingulate, statistically significant for the right rostral-anterior cingulate (Cohen’s  $d=0.53$ ). We also observed differences in cortical asymmetry. \*  $P<0.05$ .



**Fig. 5.** Scatterplots with head-size corrected region-of-interest volumes for major depressive disorder patients by remission status and healthy control subjects. Each symbol represents one individual. Line represents a best fit linear trend for the data. Non-remission (i.e., fully remitted, partially remitted, non-remitted) is associated with head-size corrected (% ICV) smaller volumes of the hippocampus, rostral-middle frontal gyrus, and right rostral-ACC, but not caudal-ACC and left rostral-ACC. Abbreviations: HC, healthy control subjects; MDD, major depressive disorder; F-R, full-remitters; P-R, partial-remitters; N-R, non-remitters. \* $P < 0.05$ , + $P < 0.10$

Table 1

Sample characteristics for major depressive disorder patients and healthy control subjects

| Variable                                                                                                               | MDD by remission status |                         |                      |                         | Healthy controls subjects (n=15) | Statistical Analyses                               |                                                      |
|------------------------------------------------------------------------------------------------------------------------|-------------------------|-------------------------|----------------------|-------------------------|----------------------------------|----------------------------------------------------|------------------------------------------------------|
|                                                                                                                        | Non-remitters (n=6)     | Partial-remitters (n=6) | Full-remitters (n=8) | All MDD subjects (n=20) |                                  | MDD by remission status vs. Controls               | All MDD vs. Controls                                 |
| Age, years                                                                                                             | 45.6±24.8               | 44.9±17.1               | 39.0±12.9            | 42.7±17.6               | 32.1±13.5                        | 3.24 <sup>a</sup> , <i>df</i> =3, <i>P</i> =0.36   | 98.5 <sup>c</sup> , <i>Z</i> =-1.72, <i>P</i> =0.09  |
| Sex, M/F (% F)                                                                                                         | 3/3 (50%)               | 4/2 (33%)               | 4/4 (50%)            | 11/9 (45%)              | 11/4 (27%)                       | 1.73 <sup>b</sup> , <i>df</i> =3, <i>P</i> =0.63   | 1.23 <sup>b</sup> , <i>df</i> =1, <i>P</i> =0.27     |
| Race/ethnicity                                                                                                         |                         |                         |                      |                         |                                  | 4.82 <sup>b</sup> , <i>df</i> =6, <i>P</i> =0.57   | 0.83 <sup>b</sup> , <i>df</i> =2, <i>P</i> =0.66     |
| Caucasian                                                                                                              | 6 (100%)                | 5 (83%)                 | 5 (71%)              | 16 (84%)                | 11 (85%)                         |                                                    |                                                      |
| African American                                                                                                       | 0 (0%)                  | 1 (17%)                 | 1 (14%)              | 2 (11%)                 | 2 (15%)                          |                                                    |                                                      |
| Hispanic                                                                                                               | 0 (0%)                  | 0 (0%)                  | 1 (14%)              | 1 (5%)                  | 0 (0%)                           |                                                    |                                                      |
| Indications for MRI <sup>*</sup> :                                                                                     |                         |                         |                      |                         |                                  | 14.08 <sup>b</sup> , <i>df</i> =12, <i>P</i> =0.30 | 3.21 <sup>b</sup> , <i>df</i> =4, <i>P</i> =0.52     |
| 1. Assess for transient syncope, dizziness, fainting                                                                   | 2 (33%)                 | 1 (17%)                 | 5 (63%)              | 8 (40%)                 | 10 (67%)                         |                                                    |                                                      |
| 2. Routine imaging                                                                                                     | 2 (33%)                 | 3 (50%)                 | 0 (0%)               | 5 (25%)                 | 3 (20%)                          |                                                    |                                                      |
| 3. Assess for transient aphasia, ataxia, vertigo, altered mental state, one-sided transient sensation loss or numbness | 1 (17%)                 | 1 (17%)                 | 2 (25%)              | 4 (20%)                 | 1 (7%)                           |                                                    |                                                      |
| 4. Headaches/migraines                                                                                                 | 0 (0%)                  | 1 (17%)                 | 1 (13%)              | 2 (10%)                 | 1 (7%)                           |                                                    |                                                      |
| 5. Rule out brain metastases                                                                                           | 1 (17%)                 | 0 (0%)                  | 0 (0%)               | 1 (5%)                  | 0 (0%)                           |                                                    |                                                      |
| Intracranial volume, mL                                                                                                | 1456±116                | 1475±179                | 1387±113             | 1434±135                | 1500±150                         | 4.70 <sup>a</sup> , <i>df</i> =3, <i>P</i> =0.20   | 92.0 <sup>c</sup> , <i>Z</i> =-1.93, <i>P</i> =0.053 |

Note: mean±SD or *n*(%). Abbreviations: MDD, major depressive disorder.<sup>a</sup>The Kruskal-Wallis H value.<sup>b</sup>The Pearson's  $\chi^2$  value.<sup>c</sup>The Mann-Whitney U value.<sup>\*</sup> Each subject's exam was reviewed by a clinical radiologist and concluded as "normal" or "unremarkable MRI of the brain."

**Table 2**

Image acquisition parameters for major depressive disorder patients and healthy control subjects. All magnetic resonance images for this study were coronal acquired T1-weighted images.

|                                                            | All MDD ( <i>n</i> =20) | Healthy control subjects ( <i>n</i> =15) | <i>U/X</i> <sup>2</sup> | <i>P</i> -value |
|------------------------------------------------------------|-------------------------|------------------------------------------|-------------------------|-----------------|
| <b>Scanner manufacturer, model</b>                         |                         |                                          |                         |                 |
| GE/Siemens (% GE)                                          | 15/5 (75%)              | 13/2 (87%)                               | 0.73 <sup>a</sup>       | 0.39            |
| GE, Genesis Signa/Signa Excite/Signa HDx (% Genesis Signa) | 8/4/3 (53%)             | 7/5/1 (54%)                              | 1.04 <sup>a</sup>       | 0.59            |
| Siemens, Sonata/Avanto (% Sonata)                          | 3/2 (60%)               | 1/1 (50%)                                | 0.06 <sup>a</sup>       | 0.81            |
| <b>Image resolution</b>                                    |                         |                                          |                         |                 |
| Field strength, Tesla                                      | 1.5 (1.5-1.5)           | 1.5 (1.5-1.5)                            | 150.0 <sup>b</sup>      | 1.00            |
| Slice thickness, mm                                        | 1.5 (1.5-1.5)           | 1.5 (1.5-1.5)                            | 147.0 <sup>b</sup>      | 0.86            |
| Matrix, 256×256/256×192/512×512 (% 256×256)                | 17/1/2 (85%)            | 13/1/1 (87%)                             | 0.16 <sup>a</sup>       | 0.93            |
| Pixel-size, mm <sup>2</sup>                                | 0.86 (0.78–0.94)        | 0.86 (0.78–0.94)                         | 131.5 <sup>b</sup>      | 0.53            |
| FOV, cm                                                    | 22 (21–24)              | 22 (22–24)                               | 148.0 <sup>b</sup>      | 0.94            |
| Number of slices                                           | 124 (124-124)           | 124 (124-124)                            | 139.0 <sup>b</sup>      | 0.63            |
| <b>Image contrast</b>                                      |                         |                                          |                         |                 |
| Repetition time, sec                                       | 3 (2.3–3)               | 3 (3-3)                                  | 121.0 <sup>b</sup>      | 0.25            |
| Echo time, ms                                              | 5 (4–8)                 | 6 (3–8)                                  | 136.0 <sup>b</sup>      | 0.63            |
| Inversion time, sec                                        | 1.1 (1.1-1.1)           | 1.1 (1.1-1.1)                            | 5.0 <sup>b</sup>        | 1.00            |
| Flip angle, degrees                                        | 35 (18–35)              | 35 (25–35)                               | 148.0 <sup>b</sup>      | 0.94            |
| Pixel bandwidth, Hz/pixel                                  | 125 (122–129)           | 122 (122–125)                            | 109.0 <sup>b</sup>      | 0.42            |

Values represent *n*(%) or Median(IQR). Abbreviations: MDD, major depressive disorder. FOV, field-of-view.

<sup>a</sup>The Pearson's  $\chi^2$  value.

<sup>b</sup>The Mann-Whitney U value.

Table 3

Group differences and hemispheric asymmetry in absolute and head-size corrected volumes for regions-of-interest in all MDD patients vs. control subjects.

| Absolute Volume (mL)                      | All MDD (n=20) |           |             |           | Healthy control (n=15) |           |             |           | Statistical analysis        |                      |                      |                        |                                                 |  |
|-------------------------------------------|----------------|-----------|-------------|-----------|------------------------|-----------|-------------|-----------|-----------------------------|----------------------|----------------------|------------------------|-------------------------------------------------|--|
|                                           | Mean           |           | SD          |           | Mean                   |           | SD          |           | MDD volume reduction        |                      | t <sup>a</sup>       |                        | Within-group hemispheric asymmetry <sup>b</sup> |  |
|                                           | Mean           | SD        | Mean        | SD        | Mean                   | SD        | Mean        | SD        | MDD volume reduction        | t <sup>a</sup>       | t <sup>a</sup>       | t <sub>MDD</sub>       | t <sub>HC</sub>                                 |  |
| Left hippocampus                          | 4.204          | 0.627     | 4.638       | 0.436     | 4.204                  | 0.627     | 4.638       | 0.436     | -9.36%                      | 2.29*                | NS                   | NS                     | L>R, 2.59*                                      |  |
| Right hippocampus                         | 4.207          | 0.557     | 4.540       | 0.463     | 4.207                  | 0.557     | 4.540       | 0.463     | -7.33%                      | 1.88 <sup>†</sup>    | NS                   | NS                     | L>R, 2.59*                                      |  |
| Left rostral-AAC                          | 2.307          | 0.484     | 2.443       | 0.585     | 2.307                  | 0.484     | 2.443       | 0.585     | -5.57%                      | NS                   | NS                   | L>R, 5.76****          | L>R, 2.52*                                      |  |
| Right rostral-AAC                         | 1.708          | 0.369     | 1.984       | 0.365     | 1.708                  | 0.369     | 1.984       | 0.365     | -13.91%                     | 2.21*                | NS                   | L>R, 5.76****          | L>R, 2.52*                                      |  |
| Left caudal-AAC                           | 1.849          | 0.588     | 1.891       | 0.386     | 1.849                  | 0.588     | 1.891       | 0.386     | -2.22%                      | NS                   | NS                   | R>L, 2.47*             | NS                                              |  |
| Right caudal-AAC                          | 2.206          | 0.508     | 2.127       | 0.363     | 2.206                  | 0.508     | 2.127       | 0.363     | +3.71%                      | NS                   | NS                   | R>L, 2.47*             | NS                                              |  |
| Left rmfg                                 | 14.226         | 2.253     | 16.126      | 3.220     | 14.226                 | 2.253     | 16.126      | 3.220     | -11.78%                     | 2.06*                | NS                   | R>L, 2.40*             | NS                                              |  |
| Right rmfg                                | 15.290         | 2.612     | 17.073      | 3.048     | 15.290                 | 2.612     | 17.073      | 3.048     | -10.44%                     | 1.86 <sup>†</sup>    | NS                   | R>L, 2.40*             | NS                                              |  |
| <b>Head-size corrected volume (% ICV)</b> | <b>Mean</b>    | <b>SD</b> | <b>Mean</b> | <b>SD</b> | <b>Mean</b>            | <b>SD</b> | <b>Mean</b> | <b>SD</b> | <b>MDD volume reduction</b> | <b>t<sup>a</sup></b> | <b>t<sup>a</sup></b> | <b>t<sub>MDD</sub></b> | <b>t<sub>HC</sub></b>                           |  |
| Left hippocampus                          | 0.293          | 0.037     | 0.311       | 0.029     | 0.293                  | 0.037     | 0.311       | 0.029     | -5.79%                      | NS                   | NS                   | NS                     | L>R, 2.67*                                      |  |
| Right hippocampus                         | 0.294          | 0.033     | 0.304       | 0.027     | 0.294                  | 0.033     | 0.304       | 0.027     | -3.29%                      | NS                   | NS                   | NS                     | L>R, 2.67*                                      |  |
| Left rostral-AAC                          | 0.161          | 0.029     | 0.165       | 0.043     | 0.161                  | 0.029     | 0.165       | 0.043     | -2.42%                      | NS                   | NS                   | L>R, 5.95****          | L>R, 2.58*                                      |  |
| Right rostral-AAC                         | 0.119          | 0.019     | 0.132       | 0.020     | 0.119                  | 0.019     | 0.132       | 0.020     | -9.85%                      | 2.07*                | NS                   | L>R, 5.95****          | L>R, 2.58*                                      |  |
| Left caudal-AAC                           | 0.129          | 0.040     | 0.127       | 0.031     | 0.129                  | 0.040     | 0.127       | 0.031     | +1.57%                      | NS                   | NS                   | R>L, 2.60*             | NS                                              |  |
| Right caudal-AAC                          | 0.154          | 0.034     | 0.143       | 0.027     | 0.154                  | 0.034     | 0.143       | 0.027     | +7.69%                      | NS                   | NS                   | R>L, 2.60*             | NS                                              |  |
| Left rmfg                                 | 0.992          | 0.131     | 1.078       | 0.196     | 0.992                  | 0.131     | 1.078       | 0.196     | -7.98%                      | NS                   | NS                   | R>L, 2.38*             | NS                                              |  |
| Right rmfg                                | 1.065          | 0.148     | 1.142       | 0.192     | 1.065                  | 0.148     | 1.142       | 0.192     | -6.74%                      | NS                   | NS                   | R>L, 2.38*             | NS                                              |  |

Abbreviations: MDD, major depressive disorder; HC, healthy control subjects; rostral-AAC, rostral anterior-cingulate cortex; caudal-AAC, caudal anterior-cingulate cortex; rmfg, rostral-middle frontal gyrus. L, left; R, right; NS, not significant.

<sup>†</sup>  $P < 0.10$ .

\*  $P < 0.05$ ,

\*\*\*  $P < 0.001$ .

<sup>a</sup>Independent-samples t-test, degrees-of-freedom=33, two-tailed significance.

<sup>b</sup>Paired-samples t-test, degrees-of-freedom=19 for MDD and 14 for Controls, two tailed-significance.



Relative volumes and Spearman's correlations for regions-of-interest among major depressive disorder patients by remission status and control subjects.

**Table 4**

| Region-of-interest                 | Healthy control subjects (n=15) | MDD by remission status |                         |                     | Statistical analysis |         |
|------------------------------------|---------------------------------|-------------------------|-------------------------|---------------------|----------------------|---------|
|                                    |                                 | Full-remitters (n=8)    | Partial-remitters (n=6) | Non-remitters (n=6) | Spearman's Rho*      | P-value |
| Left hippocampus                   | 0.322 (0.283-0.333)             | 0.311 (0.274-0.334)     | 0.303 (0.281-0.324)     | 0.279 (0.229-0.308) | -0.32                | 0.03    |
| Right hippocampus                  | 0.314 (0.281-0.323)             | 0.308 (0.270-0.317)     | 0.305 (0.275-0.332)     | 0.286 (0.232-0.299) | -0.28                | 0.05    |
| Left rostral-ACC                   | 0.151 (0.133-0.204)             | 0.170 (0.148-0.187)     | 0.143 (0.125-0.172)     | 0.172 (0.147-0.179) | 0.01                 | 0.48    |
| Right rostral-ACC                  | 0.135 (0.118-0.150)             | 0.112 (0.103-0.126)     | 0.117 (0.111-0.141)     | 0.116 (0.101-0.139) | -0.26                | 0.07    |
| Left caudal-ACC                    | 0.128 (0.106-0.142)             | 0.121 (0.109-0.143)     | 0.108 (0.084-0.128)     | 0.134 (-.108-0.198) | 0.02                 | 0.46    |
| Right caudal-ACC                   | 0.153 (0.127-0.161)             | 0.151 (0.134-0.210)     | 0.160 (0.154-0.166)     | 0.143 (0.113-0.167) | 0.08                 | 0.33    |
| Left rostral-middle frontal gyrus  | 1.094 (0.974-1.214)             | 0.994 (0.911-1.073)     | 0.981 (0.885-1.073)     | 0.953 (0.846-1.099) | -0.35                | 0.02    |
| Right rostral-middle frontal gyrus | 1.215 (0.962-1.299)             | 1.063 (1.032-1.157)     | 0.991 (0.841-1.052)     | 1.112 (0.931-1.263) | -0.26                | 0.07    |

Relative volumes (% ICV) represent Median (IQR).

Abbreviations: ICV, intra-cranial volume; ACC, anterior cingulate cortex.

\* Total n=35, one-tailed significance was used given our pre-specified one-sided hypothesis: Lower gray matter volume was expected in association with non-remission, i.e. whether the MDD patient is fully remitted, partially remitted, or non-remitted.