STRUCTURAL NEUROIMAGING BIOMARKERS OF MAJOR DEPRESSIVE
DISORDER AND ANTIDEPRESSANT RESPONSE

STRUCTURAL NEUROIMAGING BIOMARKERS OF MAJOR DEPRESSIVE
DISORDER AND ANTIDEPRESSANT RESPONSE

By JEE SU SUH, B.Sc. Hons.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the

Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2022) Hamilton, Ontario (Neuroscience)

TITLE: Structural Neuroimaging Biomarkers of Major Depressive Disorder and Antidepressant Response

AUTHOR: Jee Su Suh
          B.Sc. Honours in Neuroscience (University of Alberta)

SUPERVISOR: Dr. Benicio N. Frey M.D., M.Sc., Ph.D.

COMMITTEE: Dr. Luciano Minuzzi, M.D., Ph.D.; Dr. James P. Reilly, Ph.D.

NUMBER OF PAGES: xxiv, 239.

Lay Abstract

People diagnosed with major depressive disorder experience low mood, loss of pleasure and interest in rewarding activities, sleep changes, weight changes and other psychological and physical symptoms. There are many combinations of symptoms that can be considered depression, and many different kinds of treatments that act on the brain in different ways, not everyone responds to treatment in the same way. Clinical judgment is currently the main decision criteria for deciding which medications to prescribe to individuals. However, this is largely a process of trial-and-error, so we need additional biological indicators of which treatment might work the best. MRI scans can display brain structure in detail, such that we can compare the brains of individuals diagnosed with depression against those with no psychiatric disorder. We found that individuals with depression displayed a loss of brain mass in certain regions of the brain that are associated with mood processes. Hypothalamus volume, the brain region responsible for basic physiological processes such as sleep, appetite and stress response (also affected in depression), was more strongly related to certain genetic modifications in those diagnosed with depression as well. Although there are certain biological differences we can detect in the brain and blood of individuals diagnosed with depression, these differences were not associated with how people responded to first-line antidepressant treatments.

Abstract

Introduction: There is an ongoing interdisciplinary effort to identify objective biomarkers that could improve clinical treatment outcomes in major depressive disorder (MDD). MDD encompasses symptoms ranging from the affective and vegetative to the cognitive and executive domains, mainly marked by depressed mood and loss of interest/pleasure. With MDD prevalence rates on the rise, there is an increasing need to investigate the biological correlates of symptoms and response to antidepressants, which can vary widely between patients. In this thesis, we investigated the structural neuroimaging correlates of MDD and antidepressant response in living human participants using non-invasive magnetic resonance imaging (MRI). We endeavoured to describe structural brain features related to depressive symptoms and treatment response using both cross-sectional and longitudinal designs. We additionally investigated neuroimaging findings within the greater biological context of MDD by incorporating stress variables and molecular data pertaining to gene expression and epigenetics.

Results: Analyses of baseline differences in brain structure between MDD and healthy control (HC) participants revealed that the cerebral cortex tends to be thinner in frontal and temporoparietal regions in MDD, including the middle rostral frontal cortex, orbitofrontal cortex, pars opercularis and lingual gyrus with a general emphasis on the left hemisphere. Although hypothalamus volume was not shown to be significantly different between MDD and HC groups, we observed a greater extent of epigenetic functional relevance and a stronger relationship between hypothalamus volume and DNA methylation of key genes controlling the physiological stress response (*CRHBP, FKBP5*

and *NR3C1*). Generally, with the exception of a weak correlation between left hypothalamus volume and current episode duration, we did not observe any reliable associations between structural neuroimaging features and symptom severity, antidepressant response or childhood maltreatment. Short courses of antidepressant treatment ranging from weeks to a few months did not seem to affect brain structure to an extent detectable with 3T MRI.

Conclusion**:** The results suggest that there are certain structural features associated with major depressive disorder in the unmedicated state, most reliably thinner cerebral cortex in anterior frontal regions. Hypothalamus volume may additionally be linked to epigenetic characteristics to a greater extent in the disease state. We did not observe any structural features at 3T at baseline related to short-term antidepressant response.

Keywords

major depressive disorder, structural neuroimaging, cortical thickness, hypothalamus volume, epigenetics

Acknowledgements

I am sincerely grateful to my supervisor, Dr. Benicio Frey, for supporting my work with his clear academic vision, keen research instincts and positive brand of leadership. I greatly appreciated your ability to provide the clinical context and the compelling 'why' of every project whenever I found myself in the technical weeds. Your unerring positivity helped counter self-doubt, feelings of uncertainty and paper revision stress. Most of all, thank you for providing a safe environment and the freedom to pursue what I wanted.

Thank you to Dr. Luciano Minuzzi for making neuroimaging accessible and enjoyable at the beginning, when it was most intimidating, and for always lightening up meetings with your positive attitude and candid humour. Thank you to Dr. Jim Reilly for your research wisdom, cheerfulness, and insightful comments on statistical analysis. You both made committee meetings regularly fun experiences rather than stressful ones.

I was exceedingly fortunate to work with not only my immediate lab colleagues but with over a dozen collaborators from several Canadian universities within the CAN-BIND national consortium, and I thank each of them for their support and guidance. Special thanks to Drs. Stefanie Hassel, Laura Fiori, Pradeep Raamana, and all the Principal Investigators and Trainees I had the privilege of working with directly. I have also directly benefited from work carried out by many research coordinators, administrative assistants, and volunteers, to whom I extend my thanks as well.

I am grateful to my friends, who kept me going throughout my degree with their humour, affection, and intellectual energy. Special thanks to Anastasiya Slyepchenko and

Pedro Ballester for the many hours of neurotic rants and camaraderie, and to Lauren Cudney, Arela Agako and Manpreet Sehmbi for brightening my experience with your endless capacities for empathy, warmth, and fun. Thanks to the entire Frey lab, with special mention to Nirushi Kuhathasan, Ada Nexha and Dr. Maiko Schneider, for your cheerful collegiality and companionship throughout the past 5 years.

Thank you to my family, including my younger siblings Jee Hee and Jay, for all the late-night phone calls, homecoming feasts and camping adventures, not to mention the stability and optimism you lend to my life even from hundreds to thousands of miles away. Special thanks to my parents, tireless and dedicated academics that they are, for the inspiration to undertake this journey.

Thank you to Matt Kerslake for being my best friend and partner, not just during this degree but for the last decade. Despite all you had to put up with, you endured it all with your characteristically limitless hilarity, patience, and kindness.

A final thanks to the individuals who took time out of their busy lives to participate in this research, which couldn't happen without them.

Table of Contents

List of Figures

CHAPTER 2

Figure 1: Flowchart of screening records for eligibility in meta-analysis and descriptive review as per PRISMA guidelines.

Figure 2: Brain regions exhibiting significantly different cortical thickness values meta-analyzed between MDD and control groups.

A) Coronal view.

B) Horizontal view.

C) Sagittal view.

Figure 3: Brain regions exhibiting significantly different cortical thickness values meta-analyzed between medication-naïve, first-episode MDD patients and controls.

A) Coronal view.

B) Horizontal view.

C) Sagittal view.

Figure 4: Funnel plots to assess publication bias for significant clusters identified in each meta-analysis.

Figure 5: Heterogeneity maps displaying regions exhibiting significant study heterogeneity in the full group and subgroup samples of studies.

CHAPTER 5

Figure 1: Representative view of a manually segmented hypothalamus showing 8 contiguous slices in HC participant.

Figure 2: Relationships between HV and asymmetry to demographic variables.

a) Boxplots illustrating the differences between HC and MDD.

b) Boxplots illustrating the differences between males and females, grouped by diagnostic status.

c) Scatterplots illustrating correlations with age, grouped by diagnostic status.

Figure 3: Heatmaps illustrating the correlation of DNA methylation at each CpG site with gene expression for each of the candidate genes (*CRHBP, FKBP5, NR3C1*).

Figure 4: Plots illustrating the RVI scores of the 64 explanatory variables for each elastic net model, encompassing all CpG sites as well as sex and age as reference variables.

Supplemental Figure 1: Boxplots displaying the distribution of left hypothalamus volume, before and after applying ComBat harmonization procedure.

Supplemental Figure 2: Boxplots illustrating change in HV over 8 weeks of escitalopram treatment.

List of Tables

CHAPTER 2

Supplemental Table 2. Numeric and spatial agreement between and within independent raters for manual segmentation of the hypothalamus.

Supplemental Table 3. Mean differences and associated statistics of t-tests in HV measures between MDD and HC.

Supplemental Table 4. CpG sites that were significantly correlated with expression levels of their respective genes (within MDD and HC separately).

List of Abbreviations

5-HT	5-hydroxytryptamine
ACC	Anterior cingulate cortex
ACTH	Adrenocorticotropic hormone
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
ARI-16	Aripiprazole responders at 16 weeks
BA	Brodmann Area
BDI	Beck's Depression Inventory
BREF	World Health Organization's Quality of Life Assessment - BREF
CAN-BIND	Canadian Biomarker Integration Network in Depression
CANMAT	Canadian Network for Mood and Anxiety Treatments
CECA	Childhood Experience of Care and Abuse
CONSORT	Consolidated Standards of Reporting Trials
CRH	Corticotropin releasing hormone
CRHBP	CRH binding protein
CSF	Cerebrospinal fluid
CT	Cortical thickness
CWP	Cluster-wise p-values
DBS	Deep brain stimulation
dlPFC	Dorsolateral prefrontal cortex
DMN	Default mode network
DSM	Diagnostic and Statistical Manual of Mental Disorders
DVS	Desvenlafaxine succinate
ECT	Electroconvulsive therapy
EEG	Electroencephalography
ELECT-TDCS	Escitalopram, Placebo and tDCS in Depression
ELS	Early life stress
ENIGMA	Enhancing Neuroimaging Genetics through Meta-Analysis
ESC-16	Escitalopram responders at week 16
ESC-8	Escitalopram responders at week 8
FDR	False discovery rate
FKBP5	FK506 binding protein
fMRI	Functional magnetic resonance imaging
FOV	Field of view
GE	Gene expression
GLM	General linear model
GM	Gray matter
GM-TPM	Gray matter tissue probability map
GR	Glucocorticoid receptor
HAM-D	Hamilton Depression Rating Scale
HC	Healthy control

HPA	Hypothalamic-pituitary-adrenal
HV	Hypothalamus volume
ICC	Intra-class coefficient
ICV	Intracranial volume
LHV	Left hypothalamus volume
MADRS	Montgomery-Asberg Depression Rating Scale
MAE	Mean absolute error
MCU	McMaster University
MDD	Major depressive disorder
MDE	Major depressive episode
MINI	Mini International Neuropsychiatric Interview
MNI	Montreal Neurological Institute
MOOSE	Meta-analysis of Observational Studies in Epidemiology
mPFC	Medial prefrontal cortex
MRI	Magnetic resonance imaging
MSE	Mean standard error
NR-16	Non-responders at week 16
NR-8	Non-responders at week 8
OFC	Orbitofrontal cortex
PCC	Posterior cingulate cortex
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTSD	Post-traumatic stress disorder
QC	Quality control
RDoC	Research Domain Criteria
RHV	Right hypothalamus volume
RMF	Rostral middle frontal
ROI	Region of interest
RVI	Relative variable importance
SDM	Seed-based d mapping
SI	Suicidal ideation
sMRI	Structural MRI
SNRI	Serotonin norepinephrine reuptake inhibitor
SPC	Symmetrized percent change
SSRI	Selective serotonin retuptake inhibitor
TMS	Transcranial magnetic stimulation
VBM	Voxel-based morphometry
WM	White matter

Declaration of Academic Achievement

Chapter 2

J.S. Suh contributed to the study design, performed the literature search and data extraction, performed the quantitative analyses, and wrote the manuscript. M.A. Schneider assisted with the literature search, data extraction and study design. L. Minuzzi provided support for quantitative analyses. G.M. MacQueen, S.C. Strother, and S.H. Kennedy provided critical revision of the article. B.N. Frey contributed to the study design, methodology, acted as third reviewer in article selection and provided critical revision of the article.

The chapter in its entirety has been *published* in **Progress in Neuro-Psychopharmacology and Biological Psychiatry**. The final accepted manuscript version of this article is presented within this thesis.

Suh JS, Schneider MA, Minuzzi L, MacQueen GM, Strother SC, Kennedy SH, Frey BN. Cortical thickness in major depressive disorder: a systematic review and meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. Vol. 88. pp. 287-302. Copyright © 2018 Elsevier Inc. DOI: 10.1016/j.pnpbp.2018.08.008

Chapter 3

J.S. Suh performed data quality assurance protocols, image analysis and wrote the manuscript. L. Minuzzi provided neuroimaging/statistical expertise and provided clinical

treatment to participants. L.E. Cudney performed statistical analysis for clinical data and wrote corresponding sections of the manuscript. L.E. Cudney, W. Maich and M. Eltayebani assisted with recruitment and data collection and provided revision of the final manuscript. C.N Soares and B.N. Frey designed the study/methodology, provided clinical treatment to participants and revision of the final manuscript.

The chapter in its entirety has been *published* in **NeuroReport**. The final accepted manuscript version of this article is presented within this thesis.

Chapter 4

J.S. Suh performed data quality assurance protocols including manual editing of images, statistical analyses and wrote the manuscript. L. Minuzzi provided neuroimaging expertise and support. P.R. Raamana provided neuroimaging and statistical expertise, advised on quality assurance and critical revision of the manuscript. A. Davis, J. Harris, M. Zamyadi, G.L. Alders and S.R. Arnott provided support with neuroimaging data collection, methodology and quality assurance. S. Hassel provided support with project, administration, logistics and critical revision of the manuscript. G.B. Hall, R.B. Sassi, R.

Milev, R.W. Lam, G.M. MacQueen, S.C. Strother, S.H. Kennedy, and B.N Frey designed the trial and overarching study, provided support on project administration and critical revision of the manuscript.

The chapter in its entirety has been ***published*** in **NeuroImage: Clinical**. The final accepted manuscript version of this article is presented within this thesis.

Suh JS, Minuzzi L, Raamana PR, Davis A, Hall GB, Harris J, Hassel S, Zamyadi M, Arnott SR, Alders GL, Sassi RB, Milev R, Lam RW, MacQueen GM, Strother SC, Kennedy SH, Frey BN. An investigation of cortical thickness and antidepressant response in major T depressive disorder: A CAN-BIND study report. *NeuroImage: Clinical*. Vol. 25. Copyright © 2020 The Author(s). Published by Elsevier Inc. DOI: 10.1016/j.nicl.2020.102178

Chapter 5

J.S. Suh contributed to analytical design and methodology, performed manual segmentation of images, performed statistical analyses, and wrote the manuscript. L.M. Fiori provided expertise and support on molecular data analysis and wrote the corresponding sections of the Methods. M. Ali performed manual segmentation of images. K.L. Harkness provided expertise and support on childhood maltreatment data. M. Ramonas provided expert neuroradiological quality assurance on manual segmentations. L. Minuzzi provided neuroimaging expertise and support. M. Zamyadi

and S.R. Arnott provided support with neuroimaging data collection, methodology, quality assurance and critical revision of the manuscript. S. Hassel provided support with project, administration, logistics and critical revision of the manuscript. F. Farzan, J.A. Foster, D.J. Muller, S.V. Parikh, S. Rotzinger, R.B. Sassi, C.N. Soares, R. Uher, G. Turecki, R. Milev, R.W. Lam, G.M. MacQueen, S.C. Strother, S.H. Kennedy, and B.N Frey designed the trial and overarching study, provided support on project administration and critical revision of the manuscript.

The chapter in its entirety has been *published* in **Psychoneuroendocrinology**. The final accepted manuscript version of this article is presented within this thesis.

Chapter 1: Introduction

1.1 Major depressive disorder

Major depressive disorder (MDD) is a psychiatric diagnosis comprising episodes of persistent depressed mood and anhedonia that may last from weeks to months. These cardinal symptoms are often additionally accompanied by cognitive, somatic, vegetative, and dysphoric symptoms such as changes in weight or sleep, fatigue, psychomotor changes, feelings of worthlessness or guilt, difficulties concentrating or making decisions and in the extreme, suicidal thoughts and/or attempts (American Psychiatric Association, 2013). A diagnosis per the Diagnostic and Statistical Manual of Mental Disorders (DSM) requires any combination of 5 of the above symptoms, including the first two of depressed mood and anhedonia. The additional presence of specifiers in the DSM-5 including anxious distress, mixed features, peripartum onset and suicidality reflect the wide range of experiences and symptoms that are associated with the umbrella diagnosis of MDD. Clinical neuroimaging has sought to characterize MDD using non-invasive magnetic resonance imaging (MRI) for approximately 25 years and has made several advancements in pinning down general neural correlates of depression (Castanheira et al., 2019; Zhuo et al., 2019). However, progress has thus far been hindered by MR resolution, methodological variability, and the difficulty of validating neuroimaging findings in vitro (Falcone et al., 2013).

There is undoubtedly a need for this research; as of 2012, the annual and lifetime prevalence rates for MDD in Canada were around 4% and 10%, respectively (Patten et

al., 2015). In addition to being a profoundly negative subjective experience for the sufferer, it is also associated with impaired functioning and decreased quality of life (Greer et al., 2010; Habert et al., 2016). On the societal level, it is associated with a considerable level of economic burden due to loss of productivity and absenteeism, among other factors (Evans-Lacko & Knapp, 2016). As of 2021, the prevalence of MDD has risen significantly with the advent of the COVID-19 pandemic, which was associated with increased health risk and social isolation for a vast majority (COVID-19 Mental Disorders Collaborators, 2021).

Therefore, it has become more relevant than ever to discover non-invasive brain biomarkers of MDD symptoms and more importantly, response to antidepressants. MDD is associated with protracted medication trials, wherein up to 50% of people do not respond to initial first line treatment but must undergo additional courses of medication until they respond favourably (O'Leary et al., 2015). This process may take anywhere between 6 weeks up to several months and can be a period of great suffering for individuals who are already depressed. As this process is primarily guided by clinical judgment, being able to rely on neuroimaging to obtain reasonably accurate predictions of whether or not a given individual will respond to a given treatment might substantially improve outcomes. However, this pursuit is still in relatively early stages.

1.2 Neurobiology of depression: a brief overview

The biological basis of not only MDD but the normal functions that are affected in the disease state has been conjectured and debated for at least 25 years. Subcortical

regions such as the hippocampus (Sheline et al., 2019) and amygdala (Hamilton et al., 2008; Roddy et al., 2021) have been reliably implicated in volumetric and functional studies, but the exact cortical regions involved in depressive pathophysiology have yet to be identified with the same degree of certainty. Given the high likelihood that MDD is a circuit disorder (Chaudhury et al., 2015; Gold et al., 2022; Yao et al., 2022), relevant cortical regions are likely to be highly distributed.

Basic research suggests that subcortical regions such as the ventral tegmental area, nucleus accumbens, lateral habenula, hippocampus and amygdala play substantial roles in the induction and persistence of depressed mood and behaviour (Chaudhury et al., 2015; Y. Yang et al., 2018). In turn, components of the default mode network (DMN) such as the medial prefrontal cortex (mPFC), anterior cingulate cortex (ACC), posterior cingulate cortex (PCC) and precuneus are linked anatomically and functionally to the aforementioned subcortical regions and are likely to be involved with MDD pathophysiology (Vitkauskas & Mathuru, 2020). Notably, fMRI abnormalities within and between the DMN and other resting state networks have been well-replicated across independent studies (Li et al., 2022).

Emergent affective and cognitive functions are thought to be mediated by linkages between brain regions rather than arising from any single region. For instance, connections between various subregions of the PFC and the ventral tegmental area, ventral hippocampus and basolateral amygdala are hypothesized to mediate and influence aversive stimuli processing (Gunaydin et al., 2014), social memory (M. L. Phillips et al., 2019) and motivated behaviours (Sharp, 2017), respectively. The mPFC in general has

been identified as a hub structure from functional connectivity studies, further supported by structural findings (Bittar & Labonté, 2021). A decrease in synapse number has been observed in a postmortem analysis of the dorsolateral PFC (dlPFC) in subjects diagnosed with MDD (Kang et al., 2012). Overall, animal studies suggest that complex and overlapping circuitry extending throughout the brain gives rise to the affective and cognitive functions that go awry in MDD. The question is whether these basic research findings translate to detectable structural changes in vivo.

In living humans, MRI is often used as a non-invasive and well-tolerated method to interrogate the structural and functional basis of emotion and cognition, albeit at a much lower resolution and with limited experimental latitude. The study of resting state networks affected in MDD has identified the cerebellum, lingual gyrus, ACC, middle frontal gyrus, dlPFC, insula, and the amygdala as regions of interest (Dutta et al., 2014). Observational voxel-based morphometry (VBM) studies have suggested that gray matter volumes generally tend to be decreased in MDD (Zhang et al., 2018). In one study, the volume of the mPFC and ACC in the left hemisphere and various frontal and temporal regions on the right side were decreased in MDD after accounting for age, sex, and education (Kandilarova et al., 2019). Thinning of prefrontal cortical areas is thought to be associated with poorer clinical outcomes, specifically in the orbitofrontal cortex (Zhao et al., 2017), ACC (J. L. Phillips et al., 2015), left middle frontal gyrus and dlPFC (Kong et al., 2014). Interestingly, parietal regions involved in the DMN tend to exhibit increased cortical thickness rather than thinning in MDD (Chen et al., 2016; X. Yang et al., 2015). Given that each of the aforementioned prefrontal regions is part of distinct and

overlapping cortical-subcortical circuits, there may be some regions that are common to depression in general and other regions whose characteristics may vary more between individuals based on which circuit is most affected, likely correlating with differential symptom profiles. Results tend to differ between studies, especially with smaller sample sizes (<50 per group), so there is a need to study greater numbers in aggregate analyses to confirm disorder-specific neural correlates.

1.3 Environmental factors that influence MDD neurobiology

Stress is a major environmental risk factor for MDD; a prolonged and/or dysfunctional physiological stress response has been hypothesized to negatively influence neural circuits important for affective regulation, memory consolidation, motivation, and reward (Chaudhury et al., 2015; Dirven et al., 2017). This response is mainly mediated through the hypothalamic-pituitary-adrenal (HPA) axis from the hypothalamus, which releases corticotropin-releasing hormone (CRH) onto the anterior pituitary leading to release of adrenocorticotropic hormone (ACTH), with cortisol release being the endpoint of the hormonal cascade. In situations where stress is adaptive, cortisol acts to exert negative feedback on its own release by binding to glucocorticoid receptor (GR). In MDD, HPA dysfunction has been well-documented (Athira et al., 2020). Increased stress hormone levels such as cortisol have been observed in the blood, CSF and the brain (Chang et al., 2015; Holsboer, 2000). Plasma levels of CRH were observed to be increased in MDD (Lu et al., 2018) and administration of ACTH decreased hippocampal neurotrophic levels, which is linked to the depressive phenotype (Antunes et al., 2015).

Early life stress (ELS) or childhood maltreatment, in particular, have been conjectured to 'prime' an individual for developing MDD in adulthood (Athira et al., 2020; Seo et al., 2016). However, most people who experience ELS do not necessarily go on to develop MDD; in fact, it has been suggested that HPA dysfunction may have more to do with ELS than MDD (Ceruso et al., 2020). ELS can also interact with trauma in adulthood to influence stress axis responsivity and therefore susceptibility to mood disorders (Heim et al., 2002). NR3C1 (encoding GR) and FKBP5 (encoding FK506 binding protein 51, chaperone to GR), two genes directly related to glucocorticoid signalling, were found to exhibit stress-associated epigenetic changes (Park et al., 2019). For instance, the association between ELS and increased methylation at exon 1F on the NR3C1 gene is a relatively consistent finding in both central and peripheral tissues (Turecki & Meaney, 2016). However, the relationship between epigenetic profiles and brain measures in MDD is understudied; this may partly be due to the complexity of epigenetic data, where we quickly encounter the problem of multidimensionality.

The HPA axis additionally interacts with a number of other systems relevant in depression, such as the serotonergic, noradrenergic, and dopaminergic systems, to influence brain structure (Bao & Swaab, 2019). Per the principle of adaptive plasticity, chronic stress can influence dendritic density in a differential manner, where synaptic loss can occur in "stress-sensitive" areas such as the hippocampus and mPFC whereas dendritic expansion can occur in others such as the basolateral amygdala and OFC (Lupien et al., 2018; McEwen & Akil, 2020). Although the link between dysfunctional

HPA axis activity and hippocampal atrophy is relatively strong (Kim et al., 2015), changes in other brain regions associated with stress are less well-established.

One of these regions is the hypothalamus itself, where the physiological stress signal originates. CRH expression in the hypothalamus has been observed to exhibit a sex difference (Bao & Swaab, 2007), which, in conjunction with the action of sex hormones on various hypothalamic functions (Lund et al., 2004; Phumsatitpong et al., 2021), may partly explain the basis for the sex difference in the prevalence and experience of mood disorders between males and females. Interactions among the multiplicity of hormones and nuclei present in the hypothalamus have also been conjectured to be involved in producing dysphoric symptoms (Griffiths et al., 2000). In vivo studies on the hypothalamus are still in preliminary stages due to the difficulties associated with segmentation and the number of functionally distinct but morphologically indistinct sub-nuclei (Neudorfer et al., 2020; Schindler et al., 2012).

1.4 Structural neuroimaging as a tool for clinical translation

Given the uncertainty and complexity associated with the many overlapping theories of MDD pathophysiology, it is not yet feasible to predict antidepressant response from first principles. Considering the state of the field thus far, the ideal scenario utilizing neuroimaging might resemble the following, in simple terms: 1) perform an imaging scan of a patient as soon as their MDD diagnosis is clinically ascertained, 2) extract brain features via an automatic image processing pipeline, 3) insert the features into an algorithm trained on a representative population and 4) obtain a probabilistic prediction of response to a given treatment or get an estimate of the treatment type (medication, ECT,

TMS, etc.) most likely to result in a positive clinical response. If we can obtain a reasonable success rate, we could bypass what may have been an unhelpful or even harmful course of multiple treatment trials for the patient.

Structural MRI (sMRI) remains an appealing tool for discovering biomarkers for treatment outcomes in MDD, as it is non-invasive, well-tolerated with short scan times (~10 minutes for a standard 3T image) and no task is involved. Furthermore, participant motion presents less of a problem for sMRI compared to functional MRI (fMRI) paradigms. For the cortex, thickness is thought to be more sensitive/specific of neural changes than volume measures (Hogstrom et al., 2013) and there are several open-source automatic processing pipelines that are available to characterize and analyze cortical thickness, including but not limited to FreeSurfer (Dale et al., 1999; Fischl, 2012), CIVET (Zijdenbos et al., 2002), CAT in SPM 12 (Dahnke et al., 2013) and ANTS (Tustison et al., 2013).

Preliminary studies have suggested that hippocampal volume is most predictive of antidepressant response (MacQueen, 2009), followed by volume of the cingulate cortex (Chi et al., 2015), although most early studies were generally underpowered. Resting-state fMRI additionally suggests that activity in the ACC, as the hub in the DMN, was significant in predicting treatment response to escitalopram (Tian et al., 2019). The ACC was also highlighted in an EEG study, where stronger theta band connectivity between the rostral ACC and anterior insula was associated with greater response (Whitton et al., 2019). With respect to structural measures, it was observed that neither cortical thickness nor volume at baseline of any regions were predictive of SSRI response, although cortical

thickening in the rostral ACC in the first week was associated with a positive 8-week response (Bartlett et al., 2018). As the literature is still relatively sparse in this area, analyses employing larger sample sizes (>100 per group) are needed to investigate these claims further.

1.5 Aims and objectives

Our general aim was to use structural imaging measures to 1) identify brain features associated with the unmedicated state in MDD compared to HC and 2) test whether there is a structural signature prior to treatment that is associated with response to antidepressant treatment in individuals diagnosed with MDD. More specifically:

1. In Chapter 2, we perform a systematic review of existing literature on cortical thickness in MDD and quantify common neuroimaging findings via meta-analysis, including a critical appraisal to highlight limitations and recommendations for future work.

2. In Chapters 3 and 4, in two independent samples, we investigate whether cortical thickness at baseline is associated with the MDD diagnosis and/or clinical response to first-line antidepressant treatments.

3. In Chapter 5, we test whether hypothalamus volume and/or its relationship with methylation profiles of 3 HPA axis genes (CRHBP, FKBP5, NR3C1) are significantly altered in MDD.

1.6 Hypotheses

1. We expected to observe a general trend of cortical thinning in MDD across published findings, mostly in the frontal, cingulate and temporal regions.

2. For cortical thickness studies, we predicted that MDD would exhibit cortical thinning compared to HC and that antidepressant response would be associated with a distributed cortical thickness signature at baseline. Specifically, thicker cortex would be associated with future positive response, whereas thinner cortex would be associated with non-response.

3. We predicted that hypothalamus volume would differ in MDD compared to HC (in no particular direction, as previous literature was equivocal) and approached the brain-epigenetic relationship in an exploratory manner due to lack of priors.

1.8 References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub.

Antunes, M. S., Ruff, J. R., de Oliveira Espinosa, D., Piegas, M. B., de Brito, M. L. O., Rocha, K. A., de Gomes, M. G., Goes, A. T. R., Souza, L. C., Donato, F., Boeira, S. P., & Jesse, C. R. (2015). Neuropeptide Y administration reverses tricyclic antidepressant treatment-resistant depression induced by ACTH in mice. Hormones and Behavior, 73, 56–63. https://doi.org/10.1016/j.yhbeh.2015.05.018

Athira, K. V., Bandopadhyay, S., Samudrala, P. K., Naidu, V. G. M., Lahkar, M., & Chakravarty, S. (2020). An Overview of the Heterogeneity of Major Depressive Disorder: Current Knowledge and Future Prospective. Current Neuropharmacology, 18(3), 168–187. https://doi.org/10.2174/1570159X17666191001142934

Bao, A.-M., & Swaab, D. F. (2007). Gender difference in age-related number of corticotropin-releasing hormone-expressing neurons in the human hypothalamic paraventricular nucleus and the role of sex hormones. Neuroendocrinology, 85(1), 27–36. https://doi.org/10.1159/000099832

Bao, A.-M., & Swaab, D. F. (2019). The human hypothalamus in mood disorders: The HPA axis in the center. IBRO Reports, 6, 45–53. https://doi.org/10.1016/j.ibror.2018.11.008

Bartlett, E. A., DeLorenzo, C., Sharma, P., Yang, J., Zhang, M., Petkova, E., Weissman,

M., McGrath, P. J., Fava, M., Ogden, R. T., Kurian, B. T., Malchow, A., Cooper,

C. M., Trombello, J. M., McInnis, M., Adams, P., Oquendo, M. A., Pizzagalli,

D. A., Trivedi, M., & Parsey, R. V. (2018). Pretreatment and early-treatment

cortical thickness is associated with SSRI treatment response in major depressive

disorder. Neuropsychopharmacology, 43(11), 2221.

https://doi.org/10.1038/s41386-018-0122-9

Bittar, T. P., & Labonté, B. (2021). Functional Contribution of the Medial Prefrontal

Circuitry in Major Depressive Disorder and Stress-Induced Depressive-Like

Behaviors. Frontiers in Behavioral Neuroscience, 15.

https://doi.org/10.3389/fnbeh.2021.699592

Castanheira, L., Silva, C., Cheniaux, E., & Telles-Correia, D. (2019). Neuroimaging

Correlates of Depression-Implications to Clinical Practice. Frontiers in

Psychiatry, 10, 703. https://doi.org/10.3389/fpsyt.2019.00703

Ceruso, A., Martínez-Cengotitabengoa, M., Peters-Corbett, A., Diaz-Gutierrez, M. J., &

Martínez-Cengotitabengoa, M. (2020). Alterations of the HPA Axis Observed in

Patients with Major Depressive Disorder and Their Relation to Early Life Stress:

A Systematic Review. Neuropsychobiology, 79(6), 417–427.

https://doi.org/10.1159/000506484

Chang, H. S., Won, E. S., Lee, H.-Y., Ham, B.-J., Kim, Y.-G., & Lee, M.-S. (2015). The

association of proopiomelanocortin polymorphisms with the risk of major

depressive disorder and the response to antidepressants via interactions with

stressful life events. Journal of Neural Transmission, 122(1), 59–68.

https://doi.org/10.1007/s00702-014-1333-9

Chaudhury, D., Liu, H., & Han, M.-H. (2015). Neuronal correlates of depression. Cellular

and Molecular Life Sciences: CMLS, 72(24), 4825–4848.

https://doi.org/10.1007/s00018-015-2044-6

Chen, Z., Peng, W., Sun, H., Kuang, W., Li, W., Jia, Z., & Gong, Q. (2016). High-field

magnetic resonance imaging of structural alterations in first-episode, drug-naive

patients with major depressive disorder. Translational Psychiatry, 6, e942.

https://doi.org/10.1038/tp.2016.209

Chi, K. F., Korgaonkar, M., & Grieve, S. M. (2015). Imaging predictors of remission to

anti-depressant medications in major depressive disorder. Journal of Affective

Disorders, 186, 134–144. https://doi.org/10.1016/j.jad.2015.07.002

COVID-19 Mental Disorders Collaborators. (2021). Global prevalence and burden of

depressive and anxiety disorders in 204 countries and territories in 2020 due to

the COVID-19 pandemic. Lancet (London, England), 398(10312), 1700–1712.

https://doi.org/10.1016/S0140-6736(21)02143-7

Dahnke, R., Yotter, R. A., & Gaser, C. (2013). Cortical thickness and central surface

estimation. NeuroImage, 65, 336–348.

https://doi.org/10.1016/j.neuroimage.2012.09.050

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I.

Segmentation and surface reconstruction. NeuroImage, 9(2), 179–194.

https://doi.org/10.1006/nimg.1998.0395

Dirven, B. C. J., Homberg, J. R., Kozicz, T., & Henckens, M. J. a. G. (2017). Epigenetic programming of the neuroendocrine stress response by adult life stress. Journal of Molecular Endocrinology, 59(1), R11–R31. https://doi.org/10.1530/JME-17-0019

Dutta, A., McKie, S., & Deakin, J. F. W. (2014). Resting state networks in major depressive disorder. Psychiatry Research: Neuroimaging, 224(3), 139–151. https://doi.org/10.1016/j.pscychresns.2014.10.003

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: Absenteeism and presenteeism costs across eight diverse countries. Social Psychiatry and Psychiatric Epidemiology, 51(11), 1525–1537. https://doi.org/10.1007/s00127-016-1278-4

Falcone, M., Smith, R. M., Chenoweth, M. J., Bhattacharjee, A. K., Kelsoe, J. R., Tyndale, R. F., Lerman, C., & Pharmacogenomics Research Network Central Nervous System Working Group. (2013). Neuroimaging in psychiatric pharmacogenetics research: The promise and pitfalls. Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 38(12), 2327–2337. https://doi.org/10.1038/npp.2013.152

Fischl, B. (2012). FreeSurfer. NeuroImage, 62(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Gold, M. C., Yuan, S., Tirrell, E., Kronenberg, E. F., Kang, J. W. D., Hindley, L., Sherif, M., Brown, J. C., & Carpenter, L. L. (2022). Large-scale EEG neural network

changes in response to therapeutic TMS. Brain Stimulation, S1935-861X(22)00008-0. https://doi.org/10.1016/j.brs.2022.01.007

Greer, T. L., Kurian, B. T., & Trivedi, M. H. (2010). Defining and Measuring Functional. CNS Drugs, 24(4), 267–284. https://doi.org/10.2165/11530230-000000000-00000

Griffiths, J., Ravindran, A. V., Merali, Z., & Anisman, H. (2000). Dysthymia: A review of pharmacological and behavioral factors. Molecular Psychiatry, 5(3), 242–261. https://doi.org/10.1038/sj.mp.4000697

Gunaydin, L. A., Grosenick, L., Finkelstein, J. C., Kauvar, I. V., Fenno, L. E., Adhikari, A., Lammel, S., Mirzabekov, J. J., Airan, R. D., Zalocusky, K. A., Tye, K. M., Anikeeva, P., Malenka, R. C., & Deisseroth, K. (2014). Natural neural projection dynamics underlying social behavior. Cell, 157(7), 1535–1551. https://doi.org/10.1016/j.cell.2014.05.017

Habert, J., Katzman, M. A., Oluboka, O. J., McIntyre, R. S., McIntosh, D., MacQueen, G. M., Khullar, A., Milev, R. V., Kjernisted, K. D., Chokka, P. R., & Kennedy, S. H. (2016). Functional Recovery in Major Depressive Disorder: Focus on Early Optimized Treatment. The Primary Care Companion For CNS Disorders. https://doi.org/10.4088/PCC.15r01926

Hamilton, J. P., Siemer, M., & Gotlib, I. H. (2008). Amygdala volume in major depressive disorder: A meta-analysis of magnetic resonance imaging studies. Molecular Psychiatry, 13(11), 993–1000. https://doi.org/10.1038/mp.2008.57

Heim, C., Newport, D. J., Wagner, D., Wilcox, M. M., Miller, A. H., & Nemeroff, C. B. (2002). The role of early adverse experience and adulthood stress in the prediction of neuroendocrine stress reactivity in women: A multiple regression analysis. Depression and Anxiety, 15(3), 117–125. https://doi.org/10.1002/da.10015

Hogstrom, L. J., Westlye, L. T., Walhovd, K. B., & Fjell, A. M. (2013). The Structure of the Cerebral Cortex Across Adult Life: Age-Related Patterns of Surface Area, Thickness, and Gyrification. Cerebral Cortex, 23(11), 2521–2530. https://doi.org/10.1093/cercor/bhs231

Holsboer, F. (2000). The Corticosteroid Receptor Hypothesis of Depression. Neuropsychopharmacology, 23(5), 477–501. https://doi.org/10.1016/S0893-133X(00)00159-7

Kandilarova, S., Stoyanov, D., Sirakov, N., Maes, M., & Specht, K. (2019). Reduced grey matter volume in frontal and temporal areas in depression: Contributions from voxel-based morphometry study. Acta Neuropsychiatrica, 31(5), 252–257. https://doi.org/10.1017/neu.2019.20

Kang, H. J., Voleti, B., Hajszan, T., Rajkowska, G., Stockmeier, C. A., Licznerski, P., Lepack, A., Majik, M. S., Jeong, L. S., Banasr, M., Son, H., & Duman, R. S. (2012). Decreased expression of synapse-related genes and loss of synapses in major depressive disorder. Nature Medicine, 18(9), 1413–1417. https://doi.org/10.1038/nm.2886

Kim, E. J., Pellman, B., & Kim, J. J. (2015). Stress effects on the hippocampus: A critical

  review. Learning & Memory (Cold Spring Harbor, N.Y.), 22(9), 411–416.

  https://doi.org/10.1101/lm.037291.114

Kong, L., Wu, F., Tang, Y., Ren, L., Kong, D., Liu, Y., Xu, K., & Wang, F. (2014).

  Frontal-Subcortical Volumetric Deficits in Single Episode, Medication-Naïve

  Depressed Patients and the Effects of 8 Weeks Fluoxetine Treatment: A VBM-

  DARTEL Study. PLOS ONE, 9(1), e79055.

  https://doi.org/10.1371/journal.pone.0079055

Li, J., Chen, J., Kong, W., Li, X., & Hu, B. (2022). Abnormal core functional

  connectivity on the pathology of MDD and antidepressant treatment: A

  systematic review. *Journal of Affective Disorders*, *296*, 622–634.

  https://doi.org/10.1016/j.jad.2021.09.074

Lu, Y.-R., Zhang, Y., Rao, Y.-B., Chen, X., Lou, H.-F., Zhang, Y., Xie, H.-Y., Fang, P.,

  & Hu, L.-W. (2018). The changes in, and relationship between, plasma nitric

  oxide and corticotropin-releasing hormone in patients with major depressive

  disorder. Clinical and Experimental Pharmacology & Physiology, 45(1), 10–15.

  https://doi.org/10.1111/1440-1681.12826

Lund, T. D., Munson, D. J., Haldy, M. E., & Handa, R. J. (2004). Androgen inhibits,

  while oestrogen enhances, restraint-induced activation of neuropeptide neurones

  in the paraventricular nucleus of the hypothalamus. Journal of

  Neuroendocrinology, 16(3), 272–278. https://doi.org/10.1111/j.0953-

  8194.2004.01167.x

Lupien, S. J., Juster, R.-P., Raymond, C., & Marin, M.-F. (2018). The effects of chronic stress on the human brain: From neurotoxicity, to vulnerability, to opportunity. Frontiers in Neuroendocrinology, 49, 91–105. https://doi.org/10.1016/j.yfrne.2018.02.001

MacQueen, G. M. (2009). Magnetic resonance imaging and prediction of outcome in patients with major depressive disorder. Journal of Psychiatry & Neuroscience : JPN, 34(5), 343–349.

McEwen, B. S., & Akil, H. (2020). Revisiting the Stress Concept: Implications for Affective Disorders. The Journal of Neuroscience, 40(1), 12–21. https://doi.org/10.1523/JNEUROSCI.0733-19.2019

Neudorfer, C., Germann, J., Elias, G. J. B., Gramer, R., Boutet, A., & Lozano, A. M. (2020). A high-resolution in vivo magnetic resonance imaging atlas of the human hypothalamic region. Scientific Data, 7(1), 305. https://doi.org/10.1038/s41597-020-00644-6

O'Leary, O. F., Dinan, T. G., & Cryan, J. F. (2015). Faster, better, stronger: Towards new antidepressant therapeutic strategies. European Journal of Pharmacology, 753, 32–50. https://doi.org/10.1016/j.ejphar.2014.07.046

Park, C., Rosenblat, J. D., Brietzke, E., Pan, Z., Lee, Y., Cao, B., Zuckerman, H., Kalantarova, A., & McIntyre, R. S. (2019). Stress, epigenetics and depression: A systematic review. Neuroscience and Biobehavioral Reviews, 102, 139–152. https://doi.org/10.1016/j.neubiorev.2019.04.010

Patten, S. B., Williams, J. V. A., Lavorato, D. H., Fiest, K. M., Bulloch, A. G. M., & Wang, J. (2015). The prevalence of major depression is not changing. Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie, 60(1), 31–34. https://doi.org/10.1177/070674371506000107

Phillips, J. L., Batten, L. A., Tremblay, P., Aldosary, F., & Blier, P. (2015). A Prospective, Longitudinal Study of the Effect of Remission on Cortical Thickness and Hippocampal Volume in Patients with Treatment-Resistant Depression. International Journal of Neuropsychopharmacology, 18(8). https://doi.org/10.1093/ijnp/pyv037

Phillips, M. L., Robinson, H. A., & Pozzo-Miller, L. (2019). Ventral hippocampal projections to the medial prefrontal cortex regulate social memory. ELife, 8, e44182. https://doi.org/10.7554/eLife.44182

Phumsatitpong, C., Wagenmaker, E. R., & Moenter, S. M. (2021). Neuroendocrine interactions of the stress and reproductive axes. Frontiers in Neuroendocrinology, 63, 100928. https://doi.org/10.1016/j.yfrne.2021.100928

Roddy, D., Kelly, J. R., Farrell, C., Doolin, K., Roman, E., Nasa, A., Frodl, T., Harkin, A., O'Mara, S., O'Hanlon, E., & O'Keane, V. (2021). Amygdala substructure volumes in Major Depressive Disorder. NeuroImage. Clinical, 31, 102781. https://doi.org/10.1016/j.nicl.2021.102781

Schindler, S., Geyer, S., Strauß, M., Anwander, A., Hegerl, U., Turner, R., & Schönknecht, P. (2012). Structural studies of the hypothalamus and its nuclei in

mood disorders. Psychiatry Research: Neuroimaging, 201(1), 1–9.

https://doi.org/10.1016/j.pscychresns.2011.06.005

Seo, M. K., Ly, N. N., Lee, C. H., Cho, H. Y., Choi, C. M., Nhu, L. H., Lee, J. G., Lee, B.

J., Kim, G.-M., Yoon, B. J., Park, S. W., & Kim, Y. H. (2016). Early life stress

increases stress vulnerability through BDNF gene epigenetic changes in the rat

hippocampus. Neuropharmacology, 105, 388–397.

https://doi.org/10.1016/j.neuropharm.2016.02.009

Sharp, B. M. (2017). Basolateral amygdala and stress-induced hyperexcitability affect

motivated behaviors and addiction. Translational Psychiatry, 7(8), e1194.

https://doi.org/10.1038/tp.2017.161

Sheline, Y. I., Liston, C., & McEwen, B. S. (2019). Parsing the Hippocampus in

Depression: Chronic Stress, Hippocampal Volume, and Major Depressive

Disorder. Biological Psychiatry, 85(6), 436–438.

https://doi.org/10.1016/j.biopsych.2019.01.011

Tian, S., Sun, Y., Shao, J., Zhang, S., Mo, Z., Liu, X., Wang, Q., Wang, L., Zhao, P.,

Chattun, M. R., Yao, Z., Si, T., & Lu, Q. (2019). Predicting escitalopram

monotherapy response in depression: The role of anterior cingulate cortex.

Human Brain Mapping, 41(5), 1249–1260. https://doi.org/10.1002/hbm.24872

Turecki, G., & Meaney, M. J. (2016). Effects of the Social Environment and Stress on

Glucocorticoid Receptor Gene Methylation: A Systematic Review. Biological

Psychiatry, 79(2), 87–96. https://doi.org/10.1016/j.biopsych.2014.11.022

Tustison, N. J., Avants, B. B., Cook, P. A., Song, G., Das, S., Strien, N. van, Stone, J. R., & Gee, J. C. (2013). The ANTs cortical thickness processing pipeline. Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging, 8672, 126–129. https://doi.org/10.1117/12.2007128

Vitkauskas, M., & Mathuru, A. S. (2020). Total Recall: Lateral Habenula and Psychedelics in the Study of Depression and Comorbid Brain Disorders. International Journal of Molecular Sciences, 21(18), 6525. https://doi.org/10.3390/ijms21186525

Whitton, A. E., Webb, C. A., Dillon, D. G., Kayser, J., Rutherford, A., Goer, F., Fava, M., McGrath, P., Weissman, M., Parsey, R., Adams, P., Trombello, J. M., Cooper, C., Deldin, P., Oquendo, M. A., McInnis, M. G., Carmody, T., Bruder, G., Trivedi, M. H., & Pizzagalli, D. A. (2019). Pretreatment Rostral Anterior Cingulate Cortex Connectivity With Salience Network Predicts Depression Recovery: Findings From the EMBARC Randomized Clinical Trial. Biological Psychiatry, 85(10), 872–880. https://doi.org/10.1016/j.biopsych.2018.12.007

Yang, X., Wang, Y., Huang, J., Zhu, C., Liu, X., Cheung, E. F. C., Xie, G., & Chan, R. C. K. (2015). Increased prefrontal and parietal cortical thickness does not correlate with anhedonia in patients with untreated first-episode major depressive disorders. Psychiatry Research-Neuroimaging, 234(1), 144–151. https://doi.org/10.1016/j.pscychresns.2015.09.014

Yang, Y., Wang, H., Hu, J., & Hu, H. (2018). Lateral habenula in the pathophysiology of

    depression. Current Opinion in Neurobiology, 48, 90–96.

    https://doi.org/10.1016/j.conb.2017.10.024

Yao, H. K., Guet-McCreight, A., Mazza, F., Moradi Chameh, H., Prevot, T. D., Griffiths,

    J. D., Tripathy, S. J., Valiante, T. A., Sibille, E., & Hay, E. (2022). Reduced

    inhibition in depression impairs stimulus processing in human cortical

    microcircuits. Cell Reports, 38(2), 110232.

    https://doi.org/10.1016/j.celrep.2021.110232

Zhang, F.-F., Peng, W., Sweeney, J. A., Jia, Z.-Y., & Gong, Q.-Y. (2018). Brain structure

    alterations in depression: Psychoradiological evidence. CNS Neuroscience &

    Therapeutics, 24(11), 994–1003. https://doi.org/10.1111/cns.12835

Zhao, K., Liu, H., Yan, R., Hua, L., Chen, Y., Shi, J., Yao, Z., & Lu, Q. (2017). Altered

    patterns of association between cortical thickness and subcortical volume in

    patients with first episode major depressive disorder: A structural MRI study.

    Psychiatry Research-Neuroimaging, 260, 16–22.

    https://doi.org/10.1016/j.pscychresns.2016.12.001

Zhuo, C., Li, G., Lin, X., Jiang, D., Xu, Y., Tian, H., Wang, W., & Song, X. (2019). The

    rise and fall of MRI studies in major depressive disorder. Translational

    Psychiatry, 9(1), 335. https://doi.org/10.1038/s41398-019-0680-6

Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic "pipeline" analysis of

    3-D MRI data for clinical trials: Application to multiple sclerosis. IEEE

Transactions on Medical Imaging, 21(10), 1280–1291.

https://doi.org/10.1109/TMI.2002.806283

Chapter 2: Cortical thickness in major depressive disorder: a systematic review and meta-analysis

Jee Su Suh[a,b]; Maiko Abel Schneider[c]; Luciano Minuzzi[a,b,d]; Glenda M. MacQueen[e], Stephen C. Strother[f], Sidney H. Kennedy[g,h,i], Benicio N. Frey[a,b,d]

[a] MiNDS Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada
[b] Mood Disorders Program and Women's Health Concerns Clinic, St. Joseph's Healthcare Hamilton, ON, Canada
[c] Department of Psychiatry, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil
[d] Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada
[e] Department of Psychiatry, Mathison Centre for Mental Health Research and Education, Hotchkiss Brain Institute, Cumming School of Medicine, University of Calgary, AB
[f] Rotman Research Institute, Baycrest and Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada
[g] Canadian Biomarker Integration Network for Depression, St. Michael's Hospital, Toronto, ON, Canada
[h] Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[i] Arthur Sommer Rotenberg Chair in Suicide & Depression Studies, St. Michael's Hospital, Toronto, ON, Canada

Abstract

Neuroimaging studies assessing neurobiological differences between patients with major depressive disorder (MDD) and healthy controls (HC) are often hindered by small sample sizes and heterogeneity of the patient sample. We performed a comprehensive literature search for studies assessing cortical thickness between patient and control groups, including studies investigating treatment effects on cortical thickness. We identified 34 studies meeting criteria for the systematic review and used Seed-based d Mapping to meta-analyze 24 of those that met additional criteria. Analysis of the full sample of subjects (MDD=1073; HC=936) revealed significant thinning in the MDD group in the bilateral orbitofrontal gyrus (BA 11), left pars opercularis (BA 45) and left calcarine fissure/lingual gyrus (BA 17), as well as an area of significant thickening in the left supramarginal gyrus (BA 40). These results support other imaging modalities that report disruptions in various frontal and temporal areas in MDD and identify additional areas in all major cerebral lobes likely to be significant when parsing for biomarkers of treatment or relapse.

Keywords

"major depressive disorder"; "cortical thickness"; "magnetic resonance imaging"

2.1 Introduction

Major depressive disorder (MDD) is a heterogeneous condition, as illustrated by the various combinations of affective, cognitive and psychomotor deficits and varying levels of symptom severity that affect patients. Rates of remission with first-line treatments are problematic, estimated to be in the range of 30%-40% (Fava, 2003; Williams, 2017). Patients must often undergo multiple subsequent courses of antidepressants or augmentation strategies, which can prolong distress and functional impairment (Rush et al., 2008; Schulberg et al., 1997). MDD presents a formidable public health challenge, both for front-line clinicians as well as researchers, and justifies the current investment in identification of reliable and accurate biomarkers of the disease state that correlate to clinical measures such as symptom severity, antidepressant response and risk of treatment resistance.

In order to approach diagnosis and treatment effectively, there is a need to discern subtypes of illness (Kapur et al., 2012; Strawbridge et al., 2017) in terms of etiology, neurobiology, surrogate biomarkers or any combination of these (Arnow et al., 2015). Biomarkers could also inform risk for treatment resistance (Bennabi et al., 2015) or treatment response (Gallagher et al., 2007; Hashimoto et al., 2015), as discernable neurobiological indicators may precede symptom presentation (Aizenstein et al., 2014). With current magnetic resonance imaging (MRI) techniques, we can measure whole-brain volume (Bora et al., 2012), and further parcellate the brain into subcortical (Campbell et al., 2004) and cortical volumes (Lai et al., 2013). Cortical thickness, which is the distance between corresponding points on the pial and white matter boundaries of the neocortex,

contributes to cortical volume, along with cortical surface area and degree of gyrification (Hogstrom et al., 2013). Measured using T1-weighted structural images, it provides a better measure of the integrity of the cortical mantle in disease states (Abé et al., 2016), complementary to brain volume.

Generally, cortical thickness is determined by the number of cell bodies (neuronal and glial), extent of arborization of dendritic trees, neuronal size, their arrangement, as well as the extent of intracortical myelination (Narr et al., 2007; Seldon, 2007). Average brain thickness is 2.5-3mm but can range from 1.5-4.5mm depending on the region. Group comparisons of cortical thickness between MDD patients and healthy controls (HC) suggest important neurobiological differences, which may ultimately inform treatment targets. Work in this area has been aided by the proliferation of powerful open-source algorithms for cortical segmentation, such as FreeSurfer (http://surfer.nmr.mgh.harvard.edu/), which uses a surface-based tessellation method and infers cortical thickness by measuring the distance between white matter and pial surfaces at every vertex (Dale et al., 1999; Fischl, 2012), enabling whole-brain and region-of-interest (ROI) analyses between groups. The Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) neuroimaging consortium recently investigated group differences in cortical thickness between MDD and controls. They applied a meta-analytic technique to synthesize results across participating sites (Schmaal et al., 2017) with a resulting sample size of over 10 000 subjects. However, we are not aware of any systematic and quantitative comparison of the published literature outside of the consortium on the topic. Our review is differentiated from the above reference by the

sources of data as well as the results, which highlight areas of interest not identified in the ENIGMA study.

The objectives of the current study are to: 1) quantitatively meta-analyze neuroimaging findings to highlight brain regions implicated in cortical thickness changes in MDD, including patterns of significant correlation to clinical variables; 2) discuss the potential utility of cortical thickness as a biomarker in MDD, including a critical appraisal of major limitations within previous studies and make recommendations for future approaches.

2.2 Methods and Materials

Standard guidelines for systematic reviews and meta-analyses such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2010) and Guidelines for Meta-analysis of Observational Studies in Epidemiology (MOOSE) (Stroup et al., 2000) were followed. The protocol for this study was registered with the International Prospective Register of Systematic Reviews (PROSPERO; www.crd.york.ac.uk/) under the registration number CRD42017073427.

2.2.1 Literature search and study selection

Two researchers (JS and MS) independently consulted with an institution librarian to conduct the search strategy used for the literature search. The databases MEDLINE, Embase and Web of Science were searched, from time to conception to October 2017, for any original journal article, written in English, with a search strategy combining MeSH

terms and keywords. The following strategy was conceived in MEDLINE and adapted for the other databases: depression OR depressive disorder, major OR depressive disorder, treatment-resistant OR depress* AND (exp cerebral cortex AND organ size) OR cortical thickness/thickening/thinning AND brain mapping OR image processing, computer-assisted OR neuroimaging OR MRI OR magnetic resonance imaging NOT animals. RefWorks was used to export all articles retrieved by this search strategy and to de-duplicate the list.

2.2.2 Inclusion and exclusion criteria

Inclusion criteria for articles were as follows: 1) an original, peer-reviewed journal article, 2) adult human participants, 3) investigated changes in cortical thickness using neuroimaging, 4) used a clinically depressed sample confirmed through diagnostic interview, 5) was in English and 6) was a case-control or longitudinal cohort study. Additional inclusion criteria for being included in the meta-analysis was a) the reporting of standard space coordinates for maxima of significant voxel clusters and b) employing a whole-brain between-group analysis. Studies that did not meet the above criteria or met the following exclusion criteria were not included in the review: 1) any subjects in the primary depressed sample with a principal psychiatric diagnosis other than MDD (e.g. bipolar disorder, schizophrenia, etc.), 2) any subjects in the primary depressed sample had a concurrent neurological condition, 3) assessment of gray matter volume only, 4) the study was a case report, retrospective, or a case-series, or 5) did not have an available

version in English. Where authors met these criteria but had not reported stereotactic

coordinates, they were contacted and asked to provide this information.

2.2.3 Screening and data extraction

Following de-duplication, the two researchers performed independent screens of

the lists based on the above criteria: first by title, followed by abstract and then by full

text. Any disputes that could not be reconciled by discussion were resolved by an

objective third party (BNF). Data extracted from the final list of papers were as follows:

medication status, sample size, mean age of MDD sample, clinical severity score, the

statistical approach (whole brain or region-of-interest), method of statistical analysis,

scanner strength (Tesla), software used to process images, method of multiple comparison

correction, indication of manual editing for quality assurance, direction of effect in MDD

group (cortical thickening, thinning, or no difference), brain regions implicated in

significant results, any correlations to clinical variables and when available, standard

space coordinates (Montreal Neurological Institute [MNI] or Talairach) for maximal

voxels within significant clusters.

2.2.4 Protocol for meta-analysis

Coordinates of maximum value voxels of significant clusters were extracted from

each study (whole-brain analyses only) indicating direction of effect as well as

corresponding t-statistic values, when reported. Coordinate extraction for all studies was

manually carried out and checked independently by two researchers (JS and MS) to

minimize error. Text files in the required format were created for each study, with the standard space specified for the software to read, as well as a table indicating group sample sizes which specified whether the results were corrected or uncorrected. Studies with null results, comprised of empty text files, were also included. Seed-based d Mapping (SDM) was used to perform a meta-analysis by calculating the mean of the peak values across all the studies; details of this automated procedure have been described elsewhere (Radua et al., 2012). Although the software has been advertised for the use of meta-analyzing voxel-based structural studies, we justified the use of SDM for surface-based morphometric (SBM) studies given that the meta-analytic statistics are purely based on standard space coordinates, which most SBM studies provide, and an appropriate gray matter mask. Briefly, SDM first converts all coordinates to the same standard space (MNI). A gray matter mask is created for the peak values in each study by generating the corresponding clusters using a specialized Gaussian kernel. SDM then calculates the meta-analytic mean of the peak values of the generated clusters, ultimately producing a mean SDM map. The recommended threshold for significance (p=0.005) was used, based on the results of a previous study that tested various thresholds and found that p=0.005 acceptably balanced the specificity and sensitivity of the results when compared to the results of a gold standard analysis (Radua et al., 2012). The extent threshold was set at a minimum cluster size of 50 voxels to reduce noise in the final results. The software generates the peak MNI coordinates of significant clusters from the meta-analysis as well as corresponding Broadmann areas (BA). Standard space coordinates were used to find

corresponding labels in the FreeSurfer Desikan-Killiany atlas, facilitating comparison across other surface-based results in the literature.

Using SDM tools, Egger test statistics and funnel plots were generated for each significant peak to examine potential publication bias, and a visual inspection of heterogeneity was conducted using Q-H maps of inter-study heterogeneity generated for each analysis. A jackknife sensitivity analysis was then completed, which involves repeating the mean calculation while leaving out one study each time; these results for each cluster were assessed for replicability. This protocol was applied to the entire set of studies extracted for the analysis, and again for a subgroup analysis examining first-episode medication-naïve samples to isolate cortical thickness effects in the early course of MDD without the potential influence of medications.

2.3 Results

In total, 34 studies met criteria to be reviewed, with an inter-rater reliability score of 92% during screening (Figure 1). 30 of the 34 were cross-sectional designs, and 4 were longitudinal studies that assessed cortical thickness change following treatment. A summary of demographic information of the samples in each study can be found in Table 1, and details pertaining to neuroimaging acquisition parameters and data analysis are presented in Table 2. For the purpose of comparability, the cross-sectional studies have been further divided into whole-brain and ROI-based analyses for descriptive review.

2.3.1 Meta-analysis results for whole-brain, cross-section studies

A total of 24 whole-brain studies were included for meta-analysis using SDM (10 from the full list were not able to be used as a result of either lacking coordinates or not employing a whole-brain approach), 5 of which were included to account for null results (Han et al., 2014, Lan et al., 2015, Pirnia et al., 2016, Taylor et al., 2015, van Eijndhoven et al., 2016). The sample size for the overall analysis contained 1073 MDD patients and 936 healthy controls. There was cortical thinning in participants with MDD in the bilateral medial orbitofrontal cortex (OFC) (BA 11; p=0.00020), left pars opercularis (BA 45; p=0.00065) and left calcarine fissure/lingual gyrus (BA 17; p=0.00037). The left supramarginal gyrus (BA 40; p=0.0015) was thicker in MDD as compared to controls (Figure 2, Table 3).

There were 6 studies of first-episode medication-naïve MDD comprising 186 MDD patients and 193 healthy controls. There was cortical thinning in the left fusiform gyrus (BA 20; p=0.00049), bilateral orbitofrontal gyri (BA 47; left hem. p= 0.00071; right hem. p=0.00049) and right middle temporal gyrus (BA 21; p=0.00069) in the MDD group. Cortical thickening was seen in the right supramarginal gyrus (BA 40; p=0.00016), right medial OFC (BA 11; p=0.0030) and left inferior parietal gyrus (BA 19; p=0.0018) (Figure 3, Table 4).

2.3.2 Descriptive results for whole-brain studies

Four whole-brain studies did not report coordinates, which we were not able to obtain through correspondence, and were not included in the meta-analysis (Fonseka et

al., 2016; Han et al., 2017; Jaworska et al., 2014; Lener et al., 2016). In general,
significant results in both directions were localized to midline regions of the frontal and
temporal lobes and to a lesser extent in parietal lobes. Cortical thickening was found in
the left pars opercularis (Fonseka et al., 2016) and frontal poles (Jaworska et al., 2014).
However, Lener et al. (2016) found cortical thinning in the right rostral anterior cingulate
cortex and a trend that did not reach significance was observed in this same area in
another study (Han et al., 2017).

2.3.3 Descriptive results for region-of-interest studies

Five studies investigated a priori hypotheses for changes in cortical thickness in
specific brain regions, generally focused on frontal and temporal areas associated with
emotional regulation. Meier et al. (2016) investigated Broadmann areas 9, 10 and 11
corresponding to prefrontal areas and 24, 25 and 32 corresponding to the anterior
cingulate cortex. Among these, they observed cortical thinning in the right rostral and
dorsal anterior cingulate (BA 24, 32). In another study, Won and colleagues (2016) found
cortical thinning in bilateral medial OFC and in the right lateral OFC. Papmeyer et al.
(2015) investigated several areas in the temporal and frontal lobes as well as the anterior
cingulate and found thinning in the right fusiform gyrus and right parahippocampal gyrus.
Additionally, in their sample of MDD patients who had previously been high risk for the
disease and had developed it upon follow up, they found a significant group-by-time
interaction in the left inferior frontal gyrus and left precentral gyrus (decreased cortical
thickness in that group over time). In ROIs selected on the basis of response to

electroconvulsive therapy (ECT) in a study by Pirnia and colleagues (2016), patients displayed thinning in the fusiform gyrus and superior temporal cortex at baseline compared to HC. One study found no significant differences between groups in the following a priori selected ROIs: inferior temporal gyrus, dorsolateral prefrontal, orbitofrontal and anterior cingulate cortices (Phillips et al., 2015).

2.3.4 Descriptive results for longitudinal studies

Four studies included a longitudinal design to examine treatment effects in MDD; treatments were electroconvulsive therapy (ECT) (Pirnia et al. 2016; Sartorius et al., 2016; van Eijndhoven et al. 2016) and pharmacotherapy (Phillips et al. 2015), with three of them including a healthy control group for comparison.

Phillips et al. (2015) performed an ROI study examining changes in the dorsolateral prefrontal, orbitofrontal, anterior cingulate and inferior temporal cortices before and after treatment with varying antidepressant medications. They reported that the course of longitudinal change in cortical thickness was associated with remitter status in the rostral middle frontal gyrus, OFC and inferior temporal gyrus. The effects of ECT in MDD were assessed in three studies: Pirnia et al. (2016) observed cortical thickening in the anterior cingulate cortex bilaterally, superior temporal gyrus, temporal pole and parahippocampal gyrus. Similar results were found in van Eijndhoven et al. (2016), in particular bilateral thickening of the temporal pole, middle temporal gyrus, left insula and the left inferior temporal cortex, with responders exhibiting significantly larger thickness increases in bilateral insula than non-responders. Sartorius and colleagues (2016) also

found thickening of the right temporal pole and bilateral insula following ECT, although they did not report a correlation to remission status.

2.3.5 Descriptive results of correlations with clinical variables

24 studies investigated correlations between cortical thickness and clinical variables. Broadly, five categories of correlations with cortical thickness emerged: scores of general clinical presentation, scores of symptom dimensions, measures pertaining to vulnerability and risk, physiological/genotypical measures and measures of clinical response following treatment (Table 6), although in 8 studies there were no significant findings (Fonseka et al., 2016; Grieve et al., 2013; Meier et al., 2016; Niu et al., 2017; Papmeyer et al., 2015; van Tol et al., 2014; Wagner et al., 2012; Won et al., 2016).

General clinical severity measures displayed the most associations with thickness in areas of the parietal lobe, and to a lesser extent with areas in frontal and temporal lobes. Thickness values in orbitofrontal and other rostral frontal areas were associated mainly with various symptom dimensions. Factors of vulnerability and risk were related to thickness in parietal and frontal areas, while thickness indicators of longitudinal clinical response were located more medially in the anterior cingulate cortex and insula.

2.3.6 Assessment of bias in the meta-analysis

Egger tests were performed for each significant cluster identified in the meta-analysis and visualized using funnel plots to test for publication bias. Based on low bias indicators and corresponding non-significant p-values ($p > 0.05$ for all clusters), there was no evidence of publication bias among the studies included in either the overall group

analysis or the subgroup analysis (Figure 4, Table 5)—however, these funnel plots may not properly take into account variability in effect size across the studies unseen in the current analysis.

Heterogeneity between studies was assessed using heterogeneity maps produced by SDM, which were then thresholded with the same parameters as the statistical maps generated for mean analyses. On visual inspection, significant heterogeneity was exhibited by 3 distinct regions in the full group analysis, two of which were in close proximity to significant clusters of mean difference in this sample of studies, specifically in BA 40 and BA 11. For the subgroup analysis, between-study heterogeneity was identified in 5 clusters, although only one of these (BA 11) was located in close proximity to any clusters significant in the subgroup mean analysis.

A jackknife analysis for the overall group revealed high replicability of findings; statistical significance of all 4 clusters were retained in 23 out of 24 studies. For the subgroup analysis consisting of medication-naïve groups only, thinning of the left fusiform gyrus (BA 20) and left lateral orbitofrontal gyrus (BA 47) proved the most replicable, remaining significant in 5 of 6 studies. Two of the clusters were replicated in 3 of the 6 studies (thickening in right medial OFC, thinning in right lateral orbitofrontal gyrus in MDD), and 3 clusters were replicated only in 1 out of 6 (thickening in right supramarginal gyrus and left inferior parietal gyrus, thinning in right middle temporal gyrus in MDD) (Table 6).

2.4 Discussion

We performed a systematic review and meta-analysis of studies investigating differences in cortical thickness between MDD subjects and healthy controls, as well as changes in cortical thickness following treatment. We summarized our findings in a side-by-side comparison with those of the ENIGMA consortium (Schmaal et al., 2017) to facilitate discussion of the significant results (Table 7).

The meta-analysis found significant thinning in bilateral medial OFC (BA 11), left pars opercularis (BA 45), left calcarine fissure/lingual gyrus (BA 17) and significant cortical thickening in the left supramarginal gyrus (BA 40) in the MDD group. These results were somewhat unexpected, given the heavy emphasis in recent studies on frontal and paralimbic cortical areas such as the anterior cingulate cortex and insula as being implicated in MDD. In contrast, the analysis of medication-naïve MDD subjects revealed cortical thinning in bilateral orbitofrontal gyrus (BA 47), left fusiform gyrus (BA 20) and right middle temporal cortex (BA 21), and cortical thickening in the right supramarginal gyrus (BA 40), left inferior parietal gyrus (BA 19) and right medial OFC (BA 11), this last finding directly opposing the result of thinner bilateral medial OFC from the full-group analysis. These contrasting findings may indicate (1) a phenomenon of initial cortical thickening in the early manifestation of the disorder, counter to the suggestion that the cortical thinning observed primarily in chronic MDD may be due to the cumulative effects of neurotoxic and/or gliotoxic processes over time (Rajkowska, 2000), and/or (2) medication effects. Alternatively, cortical thickening could be associated with neurocompensatory processes in the first episode, such as via glial hypertrophy in an

attempt to combat glutamatergic neurotoxicity (Dowlati et al., 2010; Rajkowska et al., 2013; Yang et al., 2015). Another possibility is that this phenomenon may reflect an attempt to compensate for the dysregulation of limbic areas related to abnormal mood, resulting in the recruitment of other brain regions and thereby upregulating their activity (van Eijndhoven et al., 2013). In adolescents, for instance, functional hyperconnectivity was observed in several regions in the early course of illness, which may have some bearing on the observed increase in cortical thickness via increased metabolism and blood flow (Zhu et al., 2018a). This dynamic process of initial thickening followed by chronic, long-term thinning could explain our results of increased thickness of the right medial OFC in the medication-naïve, first-episode sub-analysis and of bilateral thinning in this region in the full analysis, the majority of which contain samples of patients with recurrent, chronic MDD. However, these results should be interpreted with caution, given the cross-sectional design, the minimal overlap with the overall group results, and the suboptimal replicability of 5 out of the 7 clusters. It is worth noting, however, that thinning in left fusiform gyrus and left orbitofrontal gyrus were well-replicated in this subgroup analysis, partly reflecting similar results in the analysis using the full sample of studies.

Thinning of frontal areas, particularly in the bilateral medial OFC, was a common result between ours and the ENIGMA study (Table 7) and was one of the most robust and statistically significant clusters within the overall analysis. This adds to the abundant evidence implicating frontal areas available in the literature identified using various imaging modalities. In a large meta-analysis examining general structural findings in

MDD in 143 studies, reduced volume of the OFC and gyrus rectus, which is directly adjacent to (and sometimes considered part of) the OFC, were identified (Kempton et al., 2011). Histologically, cortical thinning was reported in the rostral, middle and caudal orbitofrontal areas (BA 10, 47) in post-mortem brains of patients who had died by suicide, which was hypothesized to be related to reduced numbers and size of neurons as well as decreased glial density (Rajkowska et al., 1999). Bremner and colleagues (2002) corroborated these results by finding reduced density of neurons and glia in the gyri rectus of remitted MDD patients. The pars opercularis, which is located on the lateral surface of the frontal cortex, was implicated in our analysis as being significantly thinner in MDD and was found to exhibit volume reductions in another meta-analysis of volume morphometry in MDD (Arnone et al., 2012). The area has also been previously implicated in emotion processing, negative attention bias (Hu and Dolcos, 2017). In a recent preliminary deep brain stimulation (DBS) study, a dramatic clinical response was observed in a patient after DBS of the bilateral posterior gyrus rectus, which was correlated to strong pre-surgery connectivity between this area and the medial prefrontal cortex; they did not observe clinical response in the patients who did not display this connectivity, suggesting a neurophenotype for successful DBS treatment (Accolla et al., 2016). In addition, the medial OFC has been implicated in other functional, genetic and network studies of MDD (Kautzky et al., 2017; Long et al., 2015; Qin et al., 2014; Shimizu et al., 2015; Singh et al., 2013; Zhu et al., 2018b), suggesting abnormal function, development and/or structure of this particular region in MDD.

Recent evidence has accumulated to highlight the involvement of the occipital cortex in MDD pathophysiology. For instance, neurochemical changes in this brain region have been reported in treatment-resistant depression (Price et al., 2009), as was a negative correlation between severity of childhood maltreatment (a common risk factor for mood abnormalities in adulthood, discussed below) and inferior occipital volume (Yang et al., 2017). Abnormal tissue integrity in this region was indicated by a reduced magnetization transfer ratio in the middle occipital cortex in treatment resistant depression (Chen et al., 2016; Jia et al., 2017), corroborating graph theory results from resting state functional data showing decreased nodal centrality (connectedness) in the occipital network (Luo et al., 2015). Reductions of both functional connectivity and 5HT-1 binding in occipital cortex have also been observed (Wang et al., 2016a; Zou et al., 2016). Interestingly, Maller et al. (2014) found evidence of a phenomenon called "occipital bending," which occurs when one occipital lobe wraps around other brain regions. This was 3 times more prevalent in the MDD group than HC and was hypothesized as being due to incomplete neuronal pruning.

There is emerging evidence for an altered structural network involving the left OFC, left gyrus rectus, insula and middle occipital gyrus among other subcortical areas in treatment-naïve depressed patients (Long et al., 2015; Qin et al., 2014). More recently, a study examining genetic polymorphisms of the 5HT-1 receptor gene found binding potential differences in the gyrus rectus, inferior frontal occipital gyrus and lingual gyrus among other areas specifically in MDD patients who were GG allele carriers of the rs6295 single-nucleotide polymorphism (Kautzky et al., 2017). Taken together, our

cortical thickness results are consistent with these studies, suggesting that patterns are emerging in neuroimaging data pointing to the role of brain areas that were not commonly discussed in the classic literature as being key players in the MDD pathophysiology, suggesting new avenues for future investigation.

Our results of cortical thinning in the left fusiform gyrus in medication-naïve subjects is also consistent with the report from the ENIGMA consortium, which identified thinning of bilateral fusiform gyri in first-episode MDD (Schmaal et al., 2017). In addition, thickening of the temporal cortices were particularly found to be affected by ECT (Pirnia et al., 2016), but not always related to response status (van Eijndhoven et al., 2016). Functionally, a recent preliminary study found that treatment of refractory depression with psilocybin resulted in a decrease in cerebral blood flow in this region (Carhart-Harris et al., 2017). Other ECT studies have found normalization of lower oxygenated hemoglobin values in the frontal and temporal cortices (Hirano et al., 2017). While these results converge towards changes in temporal cortical thickness as a neural correlate in MDD and possibly treatment response, there is ambiguity in its correlation with symptom severity (Table 6).

In general, our results of significant thickness changes in the left hemisphere align with evidence suggesting its role not only in general emotion processing but more specifically with respect to MDD pathophysiology. A recent resting state study found left-sided abnormalities in spontaneous neural activity of the cortico-limbic-striatal system (implicating the prefrontal cortex, temporal cortex and limbic areas) in MDD compared to bipolar patients (Jiang et al., 2017); similarly, another functional study that

found a greater left hemispheric response to positive stimuli during general emotion processing (Balconi et al., 2015). Significant gray matter and functional changes were found exclusively in the left hemisphere in MDD patients compared to HC, including decreased gray matter in the left lingual gyrus, one of the areas of thinning identified by our analysis (Yang et al., 2015a). An effect of lateralization was also seen with respect to antidepressant response in the amygdala, where the left and right exhibited differing extents of activation as symptom severity decreased (Chen et al., 2014). Altogether, these results suggest that the left hemisphere may exhibit particular features that are affected by the pathophysiology of MDD, or perhaps that those who exhibit these irregularities of lateralization are more susceptible to developing the disorder.

To date, only 4 out of 34 studies used a longitudinal study design to examine post-treatment effects; successful treatment seems to be related to increased cortical thickness in regions generally corresponding well with between-group results of cortical differences in MDD found in the literature. In a recent transcranial magnetic stimulation (TMS) study, gray matter in the dorsolateral prefrontal cortex (PFC), left ACC, left insula, left superior temporal gyrus and right angular gyrus increased following TMS, although only ACC volume increase was correlated to clinical improvement (Lan et al., 2016). In several other studies, structural changes following various antidepressant treatments particularly have been identified in prefrontal and anterior cingulate regions, which are preserved during remission (Dusi et al., 2015). However, only one study looked at the effects of pharmacotherapy on cortical thickness in an adult population (Phillips et al., 2015) and could not isolate the effects of each different type of medication. In addition,

all of these 4 treatment studies had modest sample sizes (range: 45 – 58). In this sense,

cortical thickness studies examining treatment effects are still preliminary and lack

consensus. There is a need for larger, well-controlled treatment studies, ideally with a

placebo/control arm. Studies currently pursuing this aim with large, multi-site designs

include CAN-BIND (Lam et al., 2016), EMBARC (Trivedi et al., 2016) and ELECT-

TDCS (Brunoni et al., 2015).

      Our meta-analysis found increased thickness of the supramarginal gyrus in MDD

in both the overall and subgroup analyses. Notably, two studies (Na et al, 2016; Yang et

al., 2015) found a positive correlation between supramarginal gyrus thickness and

symptom severity scores, whereas the correlation for this measure was negative for the

left lingual gyrus (Na et al., 2016), left superior temporal and left precentral gyrus (Zorlu

et al., 2016), as might be expected given overall cortical thinning in the disorder.

Elsewhere, baseline degree centrality of the supramarginal gyrus was highly correlated

with 2-week clinical response to antidepressants, suggesting that activity of the posterior

default mode network, of which this region is a part, may have potential as a biomarker

for treatment response as well as a promising target for therapeutic action (Shen et al.,

2015). Another resting state functional study identified altered spontaneous activity in this

region in a sample of college students displaying non-clinical depressive symptoms,

which may be a risk factor for MDD later in life (Wei et al., 2015). These results indicate

that increased thickness and activity of the supramarginal gyrus could show potential as

an early indicator of disease risk and treatment response. However, this is contentious, as

other authors have reported thinning in this area in MDD cohorts (Lener et al., 2016; Tu

et al., 2012), although it must be noted that the samples in these two studies were largely composed of recurrent MDD patients. In these cases, thinning may be observed as part of an overall neurobiological signature of the transition toward recurrent, chronic MDD.

Surprisingly, only one study formally assessed the relationship between cortical thickness and previous history of childhood trauma (Fonseka et al., 2014), despite the well-established role of early trauma in a wide range of mental health outcomes in adulthood (Gilbert et al., 2009; Jansen et al., 2016; Mesquita et al., 2017; Widom et al., 2007; Whittle et al., 2017). Neurobiologically, MDD subjects with a history of childhood abuse were found to exhibit reduced gray matter density in the right OFC (Ahn et al., 2016) and a meta-analysis using voxel-based morphometry identified an association between childhood maltreatment and smaller volumes in the right OFC, superior temporal gyrus, amygdala, insula and medial temporal regions (Lim et al., 2014). A large study testing gray matter volume between 84 MDD and 84 HC subjects found a childhood maltreatment-by-diagnosis effect in bilateral prefrontal cortex (Yang et al., 2017). Considering these data, the study of the role of trauma in future cortical thickness studies is encouraged, including the type (physical, emotional, sexual, neglect), duration and possible interaction with family history (Jansen et al, 2016).

As indicated by the Q-H maps in Figure 5, there was a significant amount of inter-study heterogeneity in several regions although only three of these regions overlapped with the significant clusters identified by the meta-analysis. Another indication of heterogeneity is the fact that we were not able to replicate the commonly seen result of thinning in the insula and the ACC, as also identified by the ENIGMA group. In one

sense, some degree of heterogeneity is to be expected, given the nature of MDD itself and how results from the original studies varied by sign and location, many of which had relatively small sample sizes. SDM lends more statistical weight to studies with larger samples but likely did not overcome this limitation. Additionally, the MDD samples included here were variable not only in terms of illness progression but also by mean age, treatment resistance, medication load, depression severity and likely the number of previous major depressive episodes, in addition to other hidden variables that were not explored or reported by the original studies.

Another significant source of heterogeneity may stem from the putative existence of 'biotypes' in MDD, characterized by differing symptom combinations and patterns of functional dysconnectivity in several brain regions (Drysdale et al., 2017). This study identified two major connectivity components, one comprising of frontostriatal and orbitofrontal features correlated with anhedonia and psychomotor retardation, and the other comprising subcortical limbic features, the cingulate cortex and lateral prefrontal areas that correlated with anxiety and insomnia. Along these two dimensions, patients were distributed relatively evenly among 4 clusters based on fMRI data. This suggests that among the general MDD population, there exists heterogeneity not only in combinations of symptoms but also in the brain features that are correlated to these combinations, and such heterogeneity may potentially influence the variability in the current analysis if these functional changes extend also to anatomical alterations between subtypes.

In interpreting the results of this meta-analysis, age-related changes in cortical thickness throughout the lifespan must be considered. In the context of normal development, the cortex undergoes thinning during adolescence and young adulthood, which is likely associated with region-dependent neuronal pruning and reorganization required for network optimization (Amlien et al., 2016; Shaw et al., 2008). Alterations in the developmental trajectory of cortical thickness other than MDD may influence these results. To date, the only way to infer the cellular substrates of these changes is to use single-timepoint post-mortem methods (Rajkowska et al., 1999; Schnack et al., 2015), from which we cannot infer developmental trajectories. The role of intracortical myelin has also been speculated upon, as it may contribute to cortical "stretching," a developmental process hypothesized to optimize functional specialization resulting in cortical thinning and simultaneous increase in cortical area (Hill et al., 2010; Hogstrom et al., 2013; Seldon, 2007). In this context, cortical thinning in the absence of surface area increase could be indicative of pathophysiological processes. For instance, a finding of the ENIGMA consortium was cortical thinning in adults without change in surface area, with the opposite effect in the adolescent sample, that is, a decrease in surface area without change in thickness (Schmaal et al., 2017). Ontogenetically independent of cortical thickness (Panizzon et al., 2009), cortical area has been found to better represent cognitive measures of intelligence than thickness (Schnack et al., 2015) and it has also been shown that individuals with larger cortical areas tend to have thinner cortices (Hogstrom et al., 2013), which may influence thickness results from group comparisons. These data serve to highlight the potential utility of including cortical surface area as a

component in the MDD biomarker panel, as a complimentary physiological marker (Liu et al., 2015).

2.4.1 Strengths and limitations

Strengths of the current systematic review and meta-analysis include the fact that this is the first to quantitatively compile results across disparate studies of cortical thickness in MDD, and the resulting meta-analysis is reasonably well-powered. Although the sample of medication-naïve studies was smaller, meta-analyzing these studies precluded the common confounding factor of past medication exposure. Another strength was the use of SDM, which allows for the inclusion of both significant and null results, the ability to incorporate opposite-sign clusters in the same parametric map, as well as provision of tools for subgroup analyses and various quality control analyses. Moreover, we performed jackknife analyses, assessment of publication bias and analysis of between-study heterogeneity to test thoroughly the robustness of the results.

The summary in Table 2 indicates that there is some ambiguity around reporting quality assurance methods, particularly information regarding quality control methods such as manual correction of FreeSurfer pre-processing output, and whether or not researchers were blind to diagnostic group when editing images. In addition, it has been suggested that FreeSurfer (which was used by all studies discussed here with the exception of 4) may be better suited for whole-brain analyses rather than ROI-based comparisons compared to other surface mapping algorithms (Zhong et al., 2010), although it is still one of the most powerful and accurate surface-based analysis tools

when compared to newer algorithms (Tustison et al., 2014). The importance of manual editing was highlighted in a recent study that validated FreeSurfer output with post-mortem measurements, which found that only the analysis performed with manually edited images produced a significant correlation between generated values and histological measurements, whereas the completely automated pipeline did not (Popescu et al., 2016). In addition, Han et al. (2006) found a slight bias for thicker cortices in scanners with a field strength of 3T as compared to 1.5T; this may have had some effect on our results given that several studies used 1.5T scanners.

In terms of limitations, we only included published data, which may not encompass all data currently available on this topic. Second, due to the relatively small number of longitudinal studies and lack of readily available stereotactic coordinates, we were unable to meta-analyze changes in cortical thickness following treatment. Third, as addressed earlier, the majority of studies included had relatively modest sample sizes, which may increase uncertainty and error in meta-analytic calculations. Fourth, although SDM allows for the inclusion of statistical parametric maps which increases sensitivity of the analysis (Radua et al., 2012), they were not utilized in this analysis as per data availability, and since SDM must infer the voxel distribution centered on published coordinates, it will not be as accurate as using the raw data (Radua and Mataix-Cols, 2009). However, this is a limitation that all neuroimaging meta-analyses of this nature must contend with. Fifth, many studies did not confirm the absence of a family history of depression in the healthy controls who were compared to the MDD patients; those with a family history of MDD can be considered at risk for the disorder, and this has been

associated with a neurobiological signature (Peterson et al., 2009). In addition, MDD is associated with structural and functional deficits encompassing various regions in all four major lobes, indicating that one individual's overall neuroimaging signature in the disease state may look very different from another's when accounting for multiple regions, various brain imaging modalities and the variability in disease state/symptom presentations.

2.5 Future directions

Given the numerous interacting factors that likely contribute to MDD, multivariate approaches are becoming increasingly more common; recent and ongoing studies are employing combinations of genetic, metabolomic, proteomic, neuroimaging and clinical markers in their analyses with the hope of developing a more sensitive biomarker panel (Lam et al., 2016; Petkova et al., 2017). For instance, associating specific genotypes with neural changes has greater potential to parse out precise etiological and pathophysiological trajectories (Kostic et al., 2016; Legge et al., 2015), and as mentioned above, several multi-site studies are now underway with the aim of increasing statistical power and incorporating multiple complementary brain measures. Furthermore, despite current challenges (Kim and Na, 2018) machine learning methods nevertheless show great promise in the pursuit of developing objective diagnostic tools for personalized treatment approaches based objectively on an individual's unique neurobiological signature (Haller et al., 2014; Kambeitz et al., 2017).

2.6 Funding

2.7 Declaration of Conflicting Interests

The Authors declare no conflict of interest with regard to the content of this manuscript.

2.8 Acknowledgements

2.9 References

Abe C, Ekman C, Sellgren C, Petrovic P, Ingvar M, Landen M. 2016. Cortical thickness, volume and surface area in patients with bipolar disorder types I and II. Journal of Psychiatry & Neuroscience 41(4):240-50.

Accolla EA, Aust S, Merkl A, Schneider G, Kuehn AA, Bajbouj M, Draganski B. 2016. Deep brain stimulation of the posterior gyrus rectus region for treatment resistant depression. J Affect Disord 194:33-7.

Aizenstein HJ, Khalaf A, Walker SE, Andreescu C. 2014. Magnetic resonance imaging predictors of treatment response in late-life depression. J Geriatr Psychiatry Neurol 27(1):24-32.

Amlien IK, Fjell AM, Tamnes CK, Grydeland H, Krogsrud SK, Chaplin TA, Rosa MGP, Walhovd KB. 2016. Organizing principles of human cortical development-thickness and area from 4 to 30 years: Insights from comparative primate neuroanatomy. Cereb Cortex 26(1):257-67.

Arnone D, McIntosh AM, Ebmeier KP, Munafo MR, Anderson IM. 2012. Magnetic resonance imaging studies in unipolar depression: Systematic review and meta-regression analyses. European Neuropsychopharmacology 22(1):1-16.

Arnow BA, Blasey C, Williams LM, Palmer DM, Rekshan W, Schatzberg AF, Etkin A, Kulkarni J, Luther JF, Rush AJ. 2015. Depression subtypes in predicting antidepressant response: A report from the iSPOT-D trial. Am J Psychiatry 172(8):743-50.

Balconi M, Grippa E, Vanutelli ME. 2015. Resting lateralized activity predicts the cortical response and appraisal of emotions: An fNIRS study. Soc Cogn Affect Neurosci 10(12):1607-14.

Bennabi D, Aouizerate B, El-Hage W, Doumy O, Moliere F, Courtet P, Nieto I, Bellivier F, Bubrovsky M, Vaiva G, et al. 2015. Risk factors for treatment resistance in unipolar depression: A systematic review. J Affect Disord 171:137-41.

Bora E, Fornito A, Pantelis C, Yuecel M. 2012. Gray matter abnormalities in major depressive disorder: A meta-analysis of voxel based morphometry studies. J Affect Disord 138(1-2):9-18.

Bremner J, Vythilingam M, Vermetten E, Nazeer A, Adil J, Khan S, Staib L, Charney D. 2002. Reduced volume of orbitofrontal cortex in major depression. Biol Psychiatry 51(4):273-9.

Brown EC, Clark DL, Hassel S, MacQueen G, Ramasubbu R. 2017. Thalamocortical connectivity in major depressive disorder. J Affect Disord 217:125-31.

Brunoni AR, Sampaio-Junior B, Moffa AH, Borrione L, Nogueira BS, Marotti Aparicio LV, Veronezi B, Moreno M, Fernandes RA, Tavares D, et al. 2015. The escitalopram versus electric current therapy for treating depression clinical study (ELECT-TDCS): Rationale and study design of a non-inferiority, triple-arm, placebo-controlled clinical trial. Sao Paulo Med J 133(3):252-63.

Campbell S, Marriott M, Nahmias C, MacQueen G. 2004. Lower hippocampal volume in patients suffering from depression: A meta-analysis. Am J Psychiatry 161(4):598-607.

Canu E, Kostic M, Agosta F, Munjiza A, Ferraro PM, Pesic D, Copetti M, Peljto A, Tosevski DL, Filippi M. 2015. Brain structural abnormalities in patients with major depression with or without generalized anxiety disorder comorbidity. J Neurol 262(5):1255-65.

Carhart-Harris RL, Roseman L, Bolstridge M, Demetriou L, Pannekoek JN, Wall MB, Tanner M, Kaelen M, McGonigle J, Murphy K, et al. 2017. Psilocybin for treatment-resistant depression: FMRI-measured brain mechanisms. Sci Rep 7:13187.

Chen Y, Huang M, Hung I, Lane H, Hou C. 2014. Right and left amygdalae activation in patients with major depression receiving antidepressant treatment, as revealed by fMRI. Behav Brain Funct 10:36.

Chen Z, Peng W, Sun H, Kuang W, Li W, Jia Z, Gong Q. 2016. High-field magnetic resonance imaging of structural alterations in first-episode, drug-naive patients with major depressive disorder. Transl Psychiatr 6:e942.

Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis - I. segmentation and surface reconstruction. Neuroimage 9(2):179-94.

Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK, Lanctot KL. 2010. A meta-analysis of cytokines in major depression. Biol Psychiatry 67(5):446-57.

Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, Fetcho RN, Zebley B, Oathes DJ, Etkin A, et al. 2017. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 23(1):28-38.

Dusi N, Barlati S, Vita A, Brambilla P. 2015. Brain structural effects of antidepressant treatment in major depression. Current Neuropharmacology 13(4):458-65.

Fava M. 2003. Diagnosis and definition of treatment-resistant depression. Biol Psychiatry 53(8):649-59.

Fischl B. 2012. Freesurfer. Neuroimage 62(2):774-81.

Fonseka BA, Jaworska N, Courtright A, MacMaster FP, MacQueen GM. 2016. Cortical thickness and emotion processing in young adults with mild to moderate depression: A preliminary study. BMC Psychiatry 16:38.

Fung G, Deng Y, Zhao Q, Li Z, Qu M, Li K, Zeng Y, Jin Z, Ma Y, Yu X, et al. 2015. Distinguishing bipolar and major depressive disorders by brain structural morphometry: A pilot study. BMC Psychiatry 15:298.

Gallagher P, Robinson LJ, Gray JM, Porter RJ, Young AH. 2007. Neurocognitive function following remission in major depressive disorder: Potential objective marker of response? Aust N Z J Psychiatry 41(1):54-61.

Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S. 2009. Child maltreatment 1 burden and consequences of child maltreatment in high-income countries. Lancet 373(9657):68-81.

Grieve SM, Korgaonkar MS, Koslow SH, Gordon E, Williams LM. 2013. Widespread reductions in gray matter volume in depression. Neuroimage-Clinical 3:332-9.

Haller S, Lovblad K, Giannakopoulos P, Van De Ville D. 2014. Multivariate pattern recognition for diagnosis and prognosis in clinical neuroimaging: State of the art, current challenges and future trends. Brain Topogr 27(3):329-37.

Han K, Choi S, Jung J, Na K, Yoon H, Lee M, Ham B. 2014. Cortical thickness, cortical and subcortical volume, and white matter integrity in patients with their first episode of major depression. J Affect Disord 155:42-8.

Han K, Won E, Sim Y, Kang J, Han C, Kim Y, Kim S, Joe S, Lee M, Tae W, et al. 2017. Influence of FKBP5 polymorphism and DNA methylation on structural changes of the brain in major depressive disorder. Scientific Reports 7:42621.

Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, Busa E, Pacheco J, Albert M, Killiany R, et al. 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. Neuroimage 32(1):180-94.

Hashimoto K. 2015. Inflammatory biomarkers as differential predictors of antidepressant response. International Journal of Molecular Sciences 16(4):7796-801.

Hill J, Inder T, Neil J, Dierker D, Harwell J, Van Essen D. 2010. Similar patterns of cortical expansion during human development and evolution. Proc Natl Acad Sci U S A 107(29):13135-40.

Hirano J, Takamiya A, Yamagata B, Hotta S, Miyasaka Y, Pu S, Iwanami A, Uchida H, Mimura M. 2017. Frontal and temporal cortical functional recovery after electroconvulsive therapy for depression: A longitudinal functional near-infrared spectroscopy study. J Psychiatr Res 91:26-35.

Hogstrom LJ, Westlye LT, Walhovd KB, Fjell AM. 2013. The structure of the cerebral cortex across adult life: Age-related patterns of surface area, thickness, and gyrification. Cerebral Cortex 23(11):2521-30.

Hu Y and Dolcos S. 2017. Trait anxiety mediates the link between inferior frontal cortex volume and negative affective bias in healthy adults. Social Cognitive & Affective Neuroscience 12(5):775-82.

Jansen K, Cardoso TA, Fries GR, Branco JC, Silva RA, Kauer-Sant'Anna M, Kapczinski F, Magalhaes PVS. 2016. Childhood trauma, family history, and their association with mood disorders in early adulthood. Acta Psychiatr Scand 134(4):281-6.

Jarnum H, Eskildsen SF, Steffensen EG, Lundbye-Christensen S, Simonsen CW, Thomsen IS, Frund E, Theberge J, Larsson E. 2011. Longitudinal MRI study of cortical thickness, perfusion, and metabolite levels in major depressive disorder. Acta Psychiatr Scand 124(6):435-46.

Jaworska N, MacMaster FP, Gaxiola I, Cortese F, Goodyear B, Ramasubbu R. 2014. A preliminary study of the influence of age of onset and childhood trauma on cortical thickness in major depressive disorder. Biomed Research International :410472.

Jia Z, Peng W, Chen Z, Sun H, Zhang H, Kuang W, Huang X, Lui S, Gong Q. 2017. Magnetization transfer imaging of treatment-resistant depression. Radiology 284(2):521-9.

Jiang X, Dai X, Edmiston EK, Zhou Q, Xu K, Zhou Y, Wu F, Kong L, Wei S, Zhou Y, et al. 2017. Alteration of cortico-limbic-striatal neural system in major depressive disorder and bipolar disorder. J Affect Disord 221:297-303.

Kambeitz J, Cabral C, Sacchet MD, Gotlib IH, Zahn R, Serpa MH, Walter M, Falkai P, Koutsouleris N. 2017. Detecting neuroimaging biomarkers for depression: A meta-analysis of multivariate pattern recognition studies. Biol Psychiatry 82(5):330-8.

Kapur S, Phillips AG, Insel TR. 2012. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 17(12):1174-9.

Kautzky A, James GM, Philippe C, Baldinger-Melich P, Kraus C, Kranz GS, Vanicek T, Gryglewski G, Wadsak W, Mitterhauser M, et al. 2017. The influence of the rs6295 gene polymorphism on serotonin-1A receptor distribution investigated with PET in patients with major depression applying machine learning. Transl Psychiatr 7:e1150.

Kempton MJ, Salvador Z, Munafo MR, Geddes JR, Simmons A, Frangou S, Williams SCR. 2011. Structural neuroimaging studies in major depressive disorder meta-analysis and comparison with bipolar disorder. Arch Gen Psychiatry 68(7):675-90.

Kim Y and Na K. 2018. Application of machine learning classification for structural brain MRI in mood disorders: Critical review from a clinical perspective. Prog Neuropsychopharmacol Biol Psychiatry 80:71-80.

Kostic M, Canu E, Agosta F, Munjiza A, Novakovic I, Dobricic V, Ferraro PM, Jerkovic VM, Pekmezovic T, Tosevski DL, et al. 2016. The cumulative effect of genetic polymorphisms on depression and brain structural integrity. Hum Brain Mapp 37(6):2173-84.

Lai C. 2013. Gray matter volume in major depressive disorder: A meta-analysis of voxel-based morphometry studies. Psychiatry Res Neuroimaging 211(1):37-46.

Lam RW, Milev R, Rotzinger S, Andreazza AC, Blier P, Brenner C, Daskalakis ZJ, Dharsee M, Downar J, Evans KR, et al. 2016. Discovering biomarkers for antidepressant response: Protocol from the canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. BMC Psychiatry 16:105.

Lan MJ, Chhetry BT, Liston C, Mann JJ, Dubin M. 2016. Transcranial magnetic stimulation of left dorsolateral prefrontal cortex induces brain morphological changes in regions associated with a treatment resistant major depressive episode: An exploratory analysis. Brain Stimul 9(4):579-85.

Lan MJ, Chhetry BT, Oquendo MA, Sublette ME, Sullivan G, Mann JJ, Parsey RV. 2014. Cortical thickness differences between bipolar depression and major depressive disorder. Bipolar Disord 16(4):378-88.

Legge RM, Sendi S, Cole JH, Cohen-Woods S, Costafreda SG, Simmons A, Farmer AE, Aitchison KJ, McGuffin P, Fu CHY. 2015. Modulatory effects of brain-derived neurotrophic factor Val66Met polymorphism on prefrontal regions in major depressive disorder. British Journal of Psychiatry 206(5):379-84.

Lener MS, Kundu P, Wong E, Dewilde KE, Tang CY, Balchandani P, Murrough JW. 2016. Cortical abnormalities and association with symptom dimensions across the depressive spectrum. J Affect Disord 190:529-36.

Lim L, Radua J, Rubia K. 2014. Gray matter abnormalities in childhood maltreatment: A voxel-wise meta-analysis. Am J Psychiatry 171(8):854-63.

Liu X, Kakeda S, Watanabe K, Yoshimura R, Abe O, Ide S, Hayashi K, Katsuki A, Umeno-Nakano W, Watanabe R, et al. 2015. Relationship between the cortical thickness and serum cortisol levels in drug-naive, first-episode patients with major depressive disorder: A surface-based morphometric study. Depress Anxiety 32(9):702-8.

Long Z, Duan X, Wang Y, Liu F, Zeng L, Zhao J, Chen H. 2015. Disrupted structural connectivity network in treatment-naive depression. Prog Neuropsychopharmacol Biol Psychiatry 56:18-26.

Luo Q, Deng Z, Qin J, Wei D, Cun L, Qiu J, Hitchman G, Xie P. 2015. Frequency dependant topological alterations of intrinsic functional connectome in major depressive disorder. Sci Rep 5:9710.

Maller JJ, Thomson RHS, Rosenfeld JV, Anderson R, Daskalakis ZJ, Fitzgerald PB. 2014. Occipital bending in depression. Brain 137:1830-7.

Meier TB, Drevets WC, Wurfel BE, Ford BN, Morris HM, Victor TA, Bodurka J, Teague TK, Dantzer R, Savitz J. 2016. Relationship between neurotoxic kynurenine metabolites and reductions in right medial prefrontal cortical thickness in major depressive disorder. Brain Behavior and Immunity 53:39-48.

Mesquita CdS, Basto-Pereira M, Maia AdC. 2017. Prevalence of childhood victimization experiences in psychiatric patients: A systematic review. Int J Mental Health Addict 15(4):967-84.

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Grp. 2010. Preferred reporting

    items for systematic reviews and meta-analyses: The PRISMA statement. Int J

    Surg 8(5):336-41.

Na K.S., Won E., Kang J., Chang H.S., Yoon H.K., Tae W.S., Kim Y.K., Lee M.S., Joe

    S.H., Kim H., et al. 2016. Brain-derived neurotrophic factor promoter

    methylation and cortical thickness in recurrent major depressive disorder.

    Scientific Reports 6:21089.

Narr KL, Woods RP, Thompson PM, Szeszko P, Robinson D, Dimtcheva T, Gurbani M,

    Toga AW, Bilder RM. 2007. Relationships between IQ and regional cortical

    gray matter thickness in healthy adults. Cereb Cortex 17(9):2163-71.

Niu M, Wang Y, Jia Y, Wang J, Zhong S, Lin J, Sun Y, Zhao L, Liu X, Huang L, et al.

    2017. Common and specific abnormalities in cortical thickness in patients with

    major depressive and bipolar disorders. Ebiomedicine 16:162-71.

Olvet DM, Delaparte L, Yeh F, DeLorenzo C, McGrath PJ, Weissman MM, Adams P,

    Fava M, Deckersbach T, McInnis MG, et al. 2016. A comprehensive

    examination of white matter tracts and connectometry in major depressive

    disorder. Depress Anxiety 33(1):56-65.

Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M,

    Jacobson K, Lyons MJ, Grant MD, Franz CE, et al. 2009. Distinct genetic

    influences on cortical surface area and cortical thickness. Cerebral Cortex

    19(11):2728-35.

Papmeyer M, Giles S, Sussmann JE, Kielty S, Stewart T, Lawrie SM, Whalley HC, McIntosh AM. 2015. Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. Biol Psychiatry 78(1):58-66.

Peng D, Shi F, Li G, Fralick D, Shen T, Qiu M, Liu J, Jiang K, Shen D, Fang Y. 2015. Surface vulnerability of cerebral cortex to major depressive disorder. Plos One 10(3):e0120704.

Perlman G., Bartlett E., Delorenzo C., Weissman M., Mcgrath P., Ogden T., Jin T., Adams P., Trivedi M., Kurian B., et al. 2017. Cortical thickness is not associated with current depression in a clinical treatment study. Hum Brain Mapp (pagination):ate of Pubaton: 2017.

Petanjek Z, Judas M, Simic G, Rasin MR, Uylings HBM, Rakic P, Kostovic I. 2011. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. Proc Natl Acad Sci U S A 108(32):13281-6.

Peterson BS, Warner V, Bansal R, Zhu H, Hao X, Liu J, Durkin K, Adams PB, Wickramaratne P, Weissman MM. 2009. Cortical thinning in persons at increased familial risk for major depression. Proc Natl Acad Sci U S A 106(15):6273-8.

Petkova E, Ogden RT, Tarpey T, Ciarleglio A, Jiang B, Su Z, Carmody T, Adams P, Kraemer HC, Grannemann BD, et al. 2017. Statistical analysis plan for stage 1 EMBARC (establishing moderators and biosignatures of antidepressant response for clinical care) study. Contemp Clin Trials Commun 6:22-30.

Phillips JL, Batten LA, Tremblay P, Aldosary F, Blier P. 2015. A prospective, longitudinal study of the effect of remission on cortical thickness and hippocampal volume in patients with treatment-resistant depression. International Journal of Neuropsychopharmacology 18(8).

Pirnia T, Joshi SH, Leaver AM, Vasavada M, Njau S, Woods RP, Espinoza R, Narr KL. 2016. Electroconvulsive therapy and structural neuroplasticity in neocortical, limbic and paralimbic cortex. Translational Psychiatry 6:e832.

Popescu V, Klaver R, Versteeg A, Voorn P, Twisk JWR, Barkhof F, Geurts JJG, Vrenken H. 2016. Postmortem validation of MRI cortical volume measurements in MS. Hum Brain Mapp 37(6):2223-33.

Price RB, Shungu DC, Mao X, Nestadt P, Kelly C, Collins KA, Murrough JW, Charney DS, Mathew SJ. 2009. Amino acid neurotransmitters assessed by proton magnetic resonance spectroscopy: Relationship to treatment resistance in major depressive disorder. Biol Psychiatry 65(9):792-800.

Qin J, Wei M, Liu H, Yan R, Luo G, Yao Z, Lu Q. 2014. Abnormal brain anatomical topological organization of the cognitive-emotional and the frontoparietal circuitry in major depressive disorder. Magn Reson Med 72(5):1397-407.

Qiu L, Lui S, Kuang W, Huang X, Li J, Li J, Zhang J, Chen H, Sweeney JA, Gong Q. 2014. Regional increases of cortical thickness in untreated, first-episode major depressive disorder. Translational Psychiatry 4:e378.

Radua J, Mataix-Cols D, Phillips ML, El-Hage W, Kronhaus DM, Cardoner N, Surguladze S. 2012. A new meta-analytic method for neuroimaging studies that

combines reported peak coordinates and statistical parametric maps. Eur Psychiat 27(8):605-11.

Radua J and Mataix-Cols D. 2009. Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. British Journal of Psychiatry 195(5):393-402.

Rajkowska G. 2000. Postmortem studies in mood disorders indicate altered numbers of neurons and glial cells. Biol Psychiatry 48(8):766-77.

Rajkowska G, Miguel-Hidalgo J, Wei J, Dilley G, Pittman S, Meltzer H, Overholser J, Roth B, Stockmeier C. 1999. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. Biol Psychiatry 45(9):1085-98.

Rajkowska G and Stockmeier CA. 2013. Astrocyte pathology in major depressive disorder: Insights from human postmortem brain tissue. Curr Drug Targets 14(11):1225-36.

Rush AJ, Wisniewski SR, Warden D, Luther JF, Davis LL, Fava M, Nierenberg AA, Trivedi MH. 2008. Selecting among second-step antidepressant medication monotherapies. Arch Gen Psychiatry 65(8):870-81.

Sartorius A., Demirakca T., Bohringer A., Clemm von Hohenberg C., Aksay S.S., Bumb J.M., Kranaster L., Ende G. 2016. Electroconvulsive therapy increases temporal gray matter volume and cortical thickness. European Neuropsychopharmacology 26(3):506-17.

Schmaal L, Hibar DP, Saemann PG, Hall GB, Baune BT, Jahanshad N, Cheung JW, van Erp TGM, Bos D, Ikram MA, et al. 2017. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts

worldwide in the ENIGMA major depressive disorder working group. Mol
Psychiatry 22(6):900-9.

Schulberg H, Block M, Madonia M, Scott C, Lave J, Rodriguez E, Coulehan J. 1997. The
'usual care' of major depression in primary care practice. Arch Fam Med
6(4):334-9.

Seldon HL. 2007. Extended neocortical maturation time encompasses speciation, fatty
acid and lateralization theories of the evolution of schizophrenia and creativity.
Med Hypotheses 69(5):1085-9.

Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, Greenstein D, Clasen
L, Evans A, Rapoport JL, et al. 2008. Neurodevelopmental trajectories of the
human cerebral cortex. J Neurosci 28(14):3586-94.

Shen Y, Yao J, Jiang X, Zhang L, Xu L, Feng R, Cai L, Liu J, Wang J, Chen W. 2015.
Sub-hubs of baseline functional brain networks are related to early improvement
following two-week pharmacological therapy for major depressive disorder.
Hum Brain Mapp 36(8):2915-27.

Shimizu Y, Yoshimoto J, Toki S, Takamura M, Yoshimura S, Okamoto Y, Yamawaki S,
Doya K. 2015. Toward probabilistic diagnosis and understanding of depression
based on functional MRI data analysis with logistic group LASSO. PLoS One
10(5):e0123524.

Singh MK, Kesler SR, Hosseini SMH, Kelley RG, Amatya D, Hamilton JP, Chen MC,
Gotlib IH. 2013. Anomalous gray matter structural networks in major depressive
disorder. Biol Psychiatry 74(10):777-85.

Spati J, Hanggi J, Doerig N, Ernst J, Sambataro F, Brakowski J, Jancke L, grosse
Holtforth M, Seifritz E, Spinelli S. 2015. Prefrontal thinning affects functional
connectivity and regional homogeneity of the anterior cingulate cortex in
depression. Neuropsychopharmacology 40(7):1640-8.

Strawbridge R, Young AH, Cleare AJ. 2017. Biomarkers for depression: Recent insights,
current challenges and future prospects. Neuropsychiatr Dis Treat 13:1245-62.

Stroup D, Berlin J, Morton S, Olkin I, Williamson G, Rennie D, Moher D, Becker B, Sipe
T, Thacker S, et al. 2000. Meta-analysis of observational studies in epidemiology
- A proposal for reporting. Jama-Journal of the American Medical Association
283(15):2008-12.

Taylor WD, Boyd B, McQuoid DR, Kudra K, Saleh A, MacFall JR. 2015. Widespread
white matter but focal gray matter alterations in depressed individuals with
thoughts of death. Prog Neuropsychopharmacol Biol Psychiatry 62:22-8.

Tham MW, Woon PS, Sum MY, Lee T, Sim K. 2011. White matter abnormalities in
major depression: Evidence from post-mortem, neuroimaging and genetic
studies. J Affect Disord 132(1-2):26-36.

Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, Oquendo MA,
Bruder G, Pizzagalli D, Toups M, et al. 2016. Establishing moderators and
biosignatures of antidepressant response in clinical care (EMBARC): Rationale
and design. J Psychiatr Res 78:11-23.

Truong W., Minuzzi L., Soares C.N., Frey B.N., Evans A.C., MacQueen G.M., Hall GBC. 2013. Changes in cortical thickness across the lifespan in major depressive disorder. Psychiatry Research - Neuroimaging 214(3):204-11.

Tu P, Chen L, Hsieh J, Bai Y, Li C, Su T. 2012. Regional cortical thinning in patients with major depressive disorder: A surface-based morphometry study. Psychiatry Research-Neuroimaging 202(3):206-13.

Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, Kandel BM, van Strien N, Stone JR, Gee JC, et al. 2014. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 99:166-79.

van Eijndhoven P, Mulders P, Kwekkeboom L, van Oostrom I, van Beek M, Janzing J, Schene A, Tendolkar I. 2016. Bilateral ECT induces bilateral increases in regional cortical thickness. Translational Psychiatry 6:e874.

van Eijndhoven P, van Wingen G, Katzenbauer M, Groen W, Tepest R, Fernandez G, Buitelaar J, Tendolkar I. 2013. Paralimbic cortical thickness in first-episode depression: Evidence for trait-related differences in mood regulation. Am J Psychiatry 170(12):1477-86.

van Tol M-, Li M, Metzger CD, Hailla N, Horn DI, Li W, Heinze HJ, Bogerts B, Steiner J, He H, et al. 2014. Local cortical thinning links to resting-state disconnectivity in major depressive disorder. Psychol Med 44(10):2053-65.

Wagner G, Schultz CC, Koch K, Schachtzabel C, Sauer H, Schlosser RG. 2012. Prefrontal cortical thickness in depressed patients with high-risk for suicidal behavior. J Psychiatr Res 46(11):1449-55.

Wang L, Zhou C, Zhu D, Wang X, Fang L, Zhong J, Mao Q, Sun L, Gong X, Xia J, et al. 2016. Serotonin-1A receptor alterations in depression: A meta-analysis of molecular imaging studies. BMC Psychiatry 16:319.

Wei X, Shen H, Ren J, Liu W, Yang R, Liu J, Wu H, Xu X, Lai L, Hu J, et al. 2015. Alteration of spontaneous neuronal activity in young adults with non-clinical depressive symptoms. Psychiatry Res Neuroimaging 233(1):36-42.

Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, Charlson FJ, Norman RE, Flaxman AD, Johns N, et al. 2013. Global burden of disease attributable to mental and substance use disorders: Findings from the global burden of disease study 2010. Lancet 382(9904):1575-86.

Whittle S, Simmons JG, Hendriksma S, Vijayakumar N, Byrne ML, Dennison M, Allen NB. 2017. Childhood maltreatment, psychopathology, and the development of hippocampal subregions during adolescence. Brain Behav 7(2):e00607.

Widom CS, DuMont K, Czaja SJ. 2007. A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. Arch Gen Psychiatry 64(1):49-56.

Williams LM. 2017. Getting personalized: Brain scan biomarkers for guiding depression interventions. Am J Psychiatry 174(6):503-5.

Wolf RC, Nolte HM, Hirjak D, Hofer S, Seidl U, Depping MS, Stieltjes B, Maier-Hein K, Sambataro F, Thomann PA. 2016. Structural network changes in patients with major depression and schizophrenia treated with electroconvulsive therapy. Eur Neuropsychopharmacol 26(9):1465-74.

Won E, Choi S, Kang J, Lee M, Ham B. 2016. Regional cortical thinning of the orbitofrontal cortex in medication-naive female patients with major depressive disorder is not associated with MAOA-uVNTR polymorphism. Annals of General Psychiatry 15:26.

Yang S, Cheng Y, Mo Y, Bai Y, Shen Z, Liu F, Li N, Jiang L, Chen W, Lu Y, et al. 2017. Childhood maltreatment is associated with gray matter volume abnormalities in patients with first-episode depression. Psychiatry Res Neuroimaging 268:27-34.

Yang X, Ma X, Li M, Liu Y, Zhang J, Huang B, Zhao L, Deng W, Li T, Ma X. 2015. Anatomical and functional brain abnormalities in unmedicated major depressive disorder. Neuropsychiatric Disease and Treatment 11:2415-23.

Yang X, Wang Y, Huang J, Zhu C, Liu X, Cheung EFC, Xie G, Chan RCK. 2015. Increased prefrontal and parietal cortical thickness does not correlate with anhedonia in patients with untreated first-episode major depressive disorders. Psychiatry Research-Neuroimaging 234(1):144-51.

Zhao K., Liu H., Yan R., Hua L., Chen Y., Shi J., Lu Q., Yao Z. 2017. Cortical thickness and subcortical structure volume abnormalities in patients with major depression with and without anxious symptoms. Brain and Behavior (pagination):ate of Pubaton: 2017.

Zhao K, Liu H, Yan R, Hua L, Chen Y, Shi J, Yao Z, Lu Q. 2017. Altered patterns of association between cortical thickness and subcortical volume in patients with first episode major depressive disorder: A structural MRI study. Psychiatry Research-Neuroimaging 260:16-22.

Zhong J, Phua DYL, Qiu A. 2010. Quantitative evaluation of LDDMM, FreeSurfer, and CARET for cortical surface mapping. Neuroimage 52(1):131-41.

Zhu J, Lin X, Lin C, Zhuo C, Yu Y. 2018. Selective functional dysconnectivity of the dorsal-anterior subregion of the precuneus in drug-naive major depressive disorder. J Affect Disord 225:676-83.

Zhu Y, Quan W, Wang H, Ma Y, Yan J, Zhang H, Dong W, Yu X. 2018. Prefrontal activation during a working memory task differs between patients with unipolar and bipolar depression: A preliminary exploratory study. J Affect Disord 225:64-70.

Zorlu N, Cropley VL, Zorlu PK, Delibas DH, Adibelli ZH, Baskin EP, Esen OS, Bora E, Pantelis C. 2017. Effects of cigarette smoking on cortical thickness in major depressive disorder. J Psychiatr Res 84:1-8.

Zou K, Gao Q, Long Z, Xu F, Sun X, Chen H, Sun X. 2016. Abnormal functional connectivity density in first-episode, drug-naive adult patients with major depressive disorder. J Affect Disord 194:153-8.

2.10 Figures and Tables

Identification

- Articles identified by initial search of online databases (n = 1790)
- Following de-duplication of articles (n = 1314)

Screening

- Title screening of articles (n = 193)
  - Articles excluded based on title (- 1121)
- Abstract screening of articles (n = 53)
  - Articles excluded based on abstract (- 140)

Eligibility

- Studies eligible for review (n = 34)
  - Articles excluded based on full text (- 19)

Included

- Studies eligible for meta-analysis (n = 24)

Figure 1. **Flowchart as per PRISMA guidelines.** The process of screening records for eligibility in meta-analysis and descriptive review is displayed.
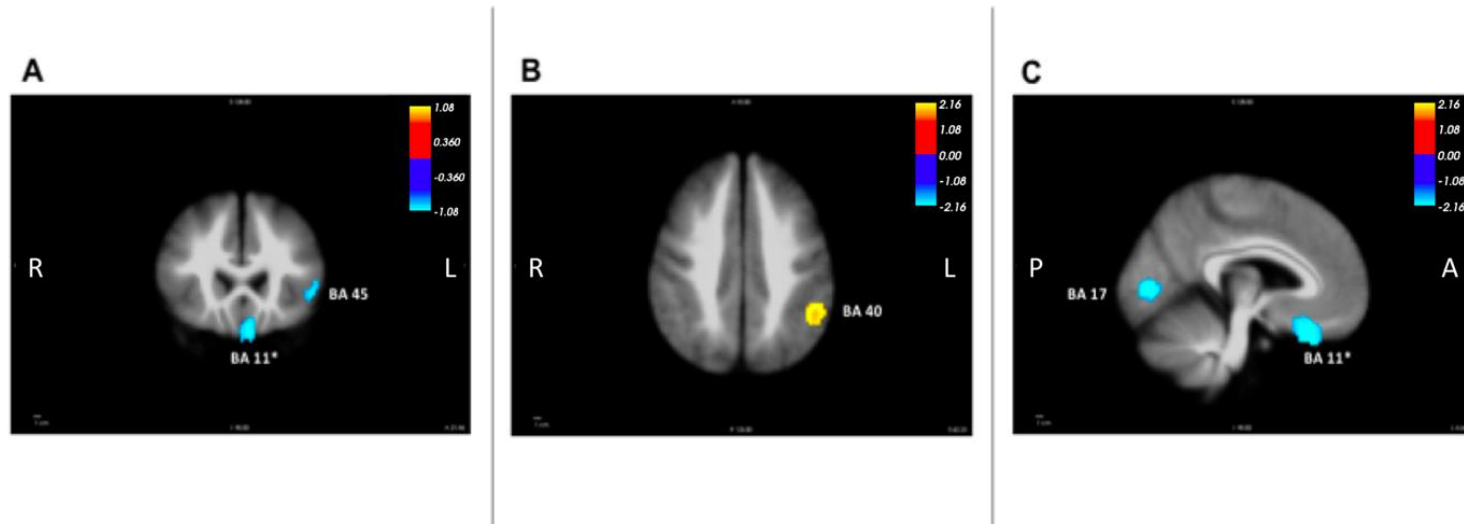
Figure 2. **Brain regions exhibiting significantly different cortical thickness values meta-analyzed between MDD and control groups.** Coronal (A), horizontal (B) and sagittal (C) cross-sectional views illustrate regions of significance identified between groups. Cooler regions indicate regions significantly thinner in the MDD group, namely in bilateral medial OFC (BA 11), left pars opercularis (BA 45) and left calcarine fissure/lingual gyrus (BA17). The warmer-coloured region indicates the region significantly thicker in the MDD group, the left supramarginal gyrus (BA 40) (p < 0.005, extent threshold = cluster size ≥ 50 voxels). The scale refers to SDM-Z scores. The asterisk (*) indicates the same cluster across cross-sectional views.
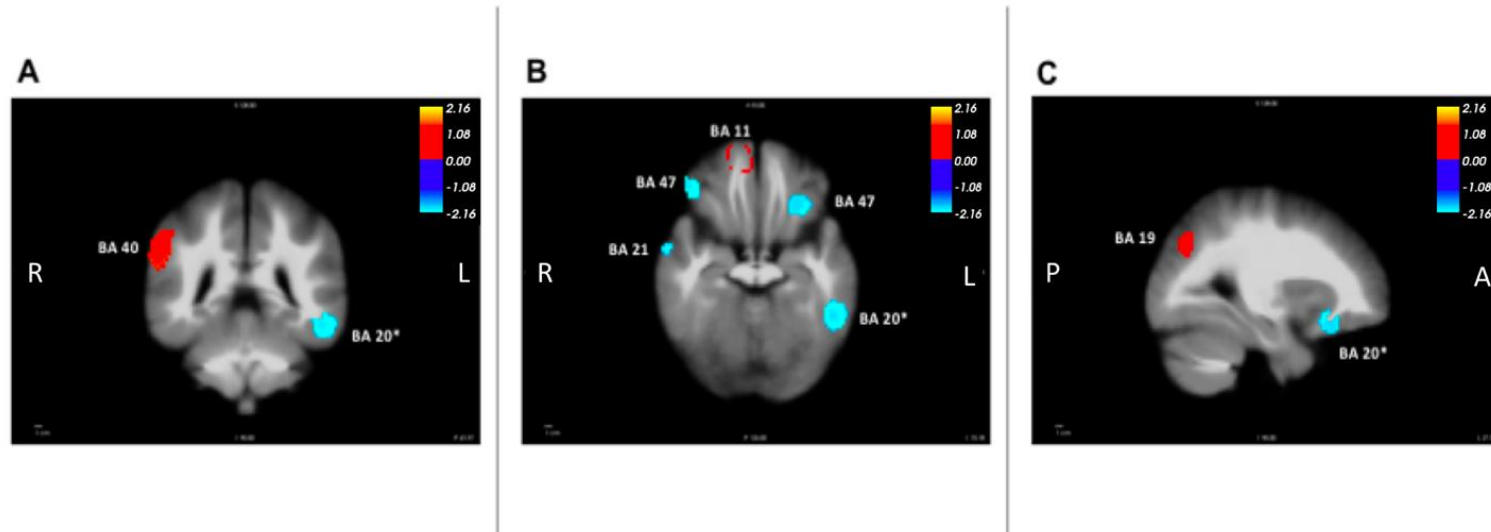
Figure 3**. Brain regions exhibiting significantly different cortical thickness values meta-analyzed between medication-naïve, first-episode MDD patients and controls.** Coronal (A), horizontal (B) and sagittal (C) cross-sectional views illustrate regions of significance identified between groups. Cooler regions indicate regions significantly thinner in the medication naïve first-episode MDD group, in left fusiform gyrus (BA 20), bilateral lateral OFC (BA 47), and right middle temporal gyrus (BA 21). The warmer colour indicates regions significantly thicker in the MDD group, specifically the right supramarginal gyrus (BA 40), right medial OFC (BA 11) and left inferior parietal gyrus (BA 19) ($p < 0.005$, extent threshold = cluster size ≥ 50 voxels). The scale refers to SDM-Z scores. The asterisk (*) indicates the same cluster across cross-sectional views.

Figure 4. **Funnel plots for significant clusters identified in each analysis.** Egger regression tests corresponding to each plot produced a p-value > 0.05 in all cases.
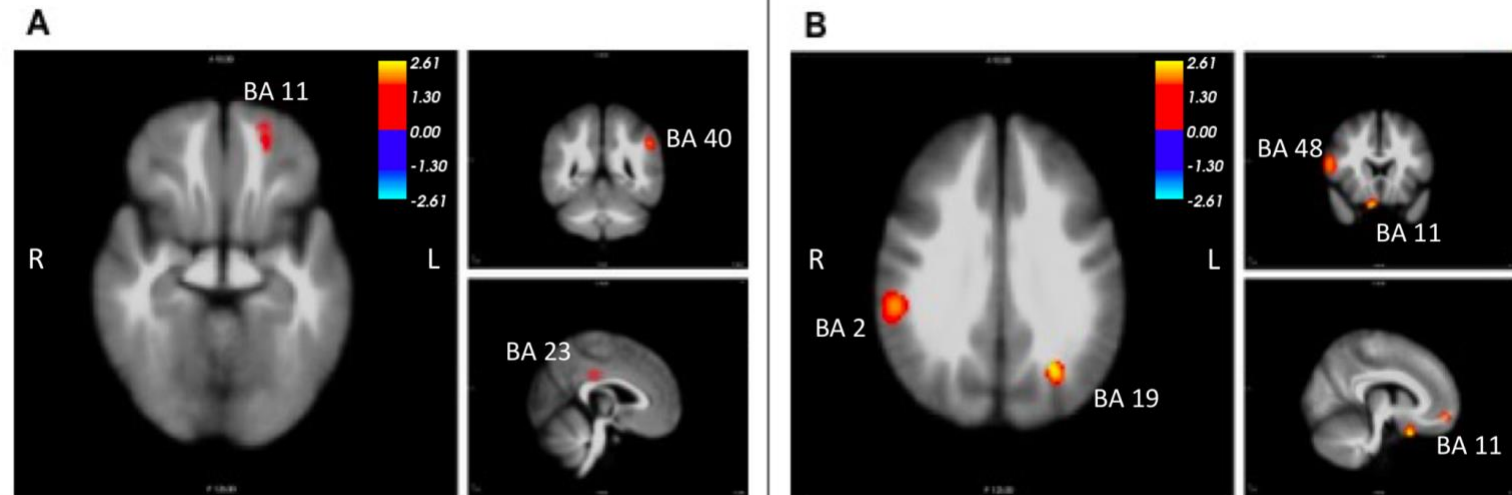
Figure 5. **Heterogeneity maps displaying regions exhibiting significant study heterogeneity in the full group and subgroup samples of studies.** The following areas were identified as having significantly variable values across the indicated sample of studies, indicated by the warmer colours. (A) In the full sample of studies, the following regions were identified as being heterogenous: left supramarginal gyrus (BA 40), left posterior cingulate (BA 23), and left orbitofrontal gyrus (BA 11). (B) In the subgroup sample, the following regions were implicated: right pars triangularis (BA 48), right inferior parietal gyrus (BA 2) and left lateral occipital gyrus (BA 19) (p < 0.005, extent threshold = cluster size ≥ 50 voxels). The scale refers to SDM-Z scores.

Table 1. **Summary of demographic information and results from included studies**

Study Name	Authors, PubYear	Sample sizes (F/M)	MDD age in years ± SD	Medication status	Mean clinical severity score	Cross-sectional results (MDD vs. HC)
**Brain structural abnormalities in patients with major depression with or without generalized anxiety disorder comorbidity	Canu et al., 2015	MDD (42/8) HC (57/14)	45.6 ± 10.2 (22-63)	All medicated	HAM-D(ns): 22.9 ± 4.7 (11-36) BDI: 32.4 ± 12.6 (7-59)	Thinning in right rostral middle frontal cortex
Cortical thickness and emotion processing in young adults with mild to moderate depression: a preliminary study	Fonseka et al., 2016	MDD (6/7) HC (7/7)	21.5 ± 1.5	>1 month medication washout	HAM-D(17): 15.3 ± 5.0	Thickening in left pars opercularis
**Distinguishing bipolar and major depressive disorders by brain structural morphometry: a pilot study	Fung et al., 2015	MDD (11/8) HC (18/11)	30.0 (sd 8.9)	8/19 unmedicated	HAM-D(17): 11.1 (sd 4.3)	Thinning in left medial OFC, left pars opercularis, left middle frontal gyri
**Widespread reductions in gray matter volume in depression	Grieve et al., 2013	MDD (54/38) HC (16/18)	33.8 ± 13.1	Not described	HAM-D(17): 21.0 ± 3.9	Thickening in left superior frontal gyrus, lateral OFC (BA 11); Thinning in left medial frontal/medial OFC (BA 25)
**Cortical thickness, cortical and subcortical volume, and white matter integrity in patients with their first episode of major depression	Han et al., 2014	MDD (15/5) HC (15/7)	42.70 ± 12.43	Medication-naïve	HAM-D(17): 19.05 ± 6.74	No significant differences
Influence of FKBP5 polymorphism and DNA methylation on structural changes of the brain in major depressive disorder	Han et al., 2017	MDD (90/24) HC (61/27)	43.51 ± 12.0	53/61 unmedicated	HAM-D(17): 14.81 ± 8.02	No significant differences (trend for thinner right ACC)
**Longitudinal MRI study of cortical thickness, perfusion, and metabolite levels in major depressive disorder	Jarnum et al., 2011	MDD (16/7) HC(13/13)	43.2 (sd 9.9)	8/15 unmedicated	HAM-D(17): 22.3 (sd 3.6)	Thinning in OFC, superior temporal lobe, insular cortex
A preliminary study of the influence of age of onset and childhood trauma on cortical	Jaworska et al., 2014	MDD (22/14) HC (10/8)	37.1 ± 11.2	3-week medication washout	HAM-D(17): 22.1 ± 4.1	Thinning in frontal pole

thickness in major depressive disorder						
**Cortical thickness differences between bipolar depression and major depressive disorder	Lan et al., 2014	MDD (32/24) HC (26/28)	36.9 (sd 12.2)	Varied lifetime medication hx	HAM-D(17): 23.8 (sd 6.0)	No significant differences
Cortical abnormalities and association with symptom dimensions across the depressive spectrum	Lener et al., 2016	MDD (29/28) HC (16/13)	40.27 ± 12.28	1-week medication washout	MADRS: 28.07 ± 6.09	Thinning in right rostral ACC
**Relationship between the cortical thickness and serum cortisol levels in drug-naive, first-episode patients with major depressive disorder: a surface-based morphometric study	Liu et al., 2015	MDD (13/17) HC (13/28)	44.9 ± 13.0 (20-67)	Medication-naïve	HAM-D(17): 21.0 ± 6.0	Thinning in left lateral OFC
Relationship between neurotoxic kynurenine metabolites and reductions in right medial prefrontal cortical thickness in major depressive disorder	Meier et al., 2016	MDD (57/16) HC (55/36)	34.2 ± 9.3	4-week medication washout	MADRS: 27.6 ± 8.8	Thinning in BA 24 and BA 32
**Brain-derived neurotrophic factor promoter methylation and cortical thickness in recurrent major depressive disorder.	Na et al., 2016	MDD (54/11) HC (50/15)	42.52 ± 11.42	33/65 medicated	HAM-D(17): 16.00 ± 8.2	Thinning in right medial OFC, right lingual, right lateral occipital, left lateral OFC, left pars triangularis, left lingual cortex
**Common and specific abnormalities in cortical thickness in patients with major depressive and bipolar disorders	Niu et al., 2017	MDD (19/17) HC (13/17)	29.08 ± 7.16	17/36 medicated	HAM-D(24): 26.67 ± 4.73	Thinning in left inferior temporal cortex
Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder	Papmeyer et al., 2015	MDD (13/7) HC (41/21)	23.33 (2.98)	4/20 medicated	HAM-D (ns): 5 (sd 12)	Thinning in right parahippocampal gyrus, right fusiform gyrus
**Surface vulnerability of cerebral cortex to major depressive disorder	Peng et al., 2015	MDD (9/7) HC (9/7)	34.43 ± 6.72	4-week medication washout	HAM-D(24): 30.88 ± 7.69	Thickening in right inferior parietal region, right paracentral gyrus, right

						transverse temporal gyrus, right posterior cingulate cortex, left superior parietal gyrus, left inferior parietal gyrus, left lateral occipital cortex; <u>Thinning</u> in right middle temporal gyrus, right pars opercularis, left pars opercularis, left rostral-middle frontal region, left precentral gyrus
**Cortical thickness is not associated with current depression in a clinical treatment study	Perlman et al., 2017	MDD (102/68) HC (29/23)	36.41 ± 12.43	All medicated	QIDS: 18.15 ± 2.89	<u>Thickening</u> in left supramarginal gyrus
A prospective, longitudinal study of the effect of remission on cortical thickness and hippocampal volume in patients with treatment-resistant depression	Phillips et al., 2015	MDD (18/8) HC (18/10)	46.0 (sd 10.4)	All medicated	MADRS: 34.6 (sd 7.0)	No significant differences at baseline
**A prospective, longitudinal study of the effect of remission on cortical thickness and hippocampal volume in patients with treatment-resistant depression	Pirnia et al., 2016	MDD (18/11) HC (16/13)	41.0 (sd 13.5)	24-48h medication washout	HAM-D(17): 26.3 (sd 4.93)	No significant differences using whole-brain analysis, thinning in the fusiform and superior temporal cortex (ROI only)
**Regional increases of cortical thickness in untreated, first-episode major depressive disorder	Qiu et al., 2014a	MDD (33/13) HC (33/13)	34.9 (sd 10.8)	Medication-naïve	HAM-D(17): 23.3 ± 5.0	<u>Thickening</u> in right frontoparietal regions (BA 9, 10), pars opercularis (BA 44), rostral middle frontal gyrus (BA 46), supramarginal gyrus (BA 40)
Electroconvulsive therapy increases temporal gray matter volume and cortical thickness	Sartorius et al., 2016	MDD (9/9) HC (18/18)	51.72 ± 13.4	All medicated	HAM-D(21): 31.8 ± 8.2	[Longitudinal analysis only]
**Prefrontal thinning affects functional connectivity and regional homogeneity of the	Spati et al., 2015	MDD (10/11) HC (20/15)	36.6 ± 12.3	16/21 medication-naïve	BDI: 26.4 ± 9.3	<u>Thinning</u> in right rostral middle frontal cortex

anterior cingulate cortex in depression.						
**Widespread white matter but focal gray matter alterations in depressed individuals with thoughts of death	Taylor et al., 2015	MDD-SI (11/10) MDD-no SI (41/12) HC (56/35)	SI: 33.5 ± 9.1 No SI: 37.5 ± 8.9	No medication in last month	MADRS: 24.7 ± 4.2 (SI), 23.3 ± 4.5 (no SI)	No significant differences
**Changes in cortical thickness across adulthood in major depressive disorder	Truong et al., 2013	MDD-EO (20/8) MDD-LO (10/11) HC (60/4)	MDD-EO: 24.47 ± 9.34 MDD-LO: 39.52 ± 8.50	Not described	HAM-D(ns): 23.36 ± 11.57 (EO), 13.29 ± 5.70 (LO)	Thinning in left dorsolateral PFC
**Regional cortical thinning in patients with major depressive disorder: A surface-based morphometry study	Tu et al., 2012	MDD (24/12) HC (22/14)	41.64 ± 12.04	All medicated	HAM-D(17): 13.25 ± 9.2	Thickening in left lateral OFC, left insula; Thinning in left superior frontal, precentral, rostral middle frontal, superior frontal, caudal middle frontal cortices; left entorhinal, middle temporal cortices; left inferior parietal cortex, left lateral occipital, left lingual cortices; right precentral, rostral middle frontal, superior frontal, caudal middle frontal cortices, lateral OFC, pars opercularis; right postcentral, supramarginal, inferior parietal cortices
**Bilateral ECT induces bilateral increases in regional cortical thickness	Van Eijndhoven et al., 2016	MDD (15/8) HC (14/8)	50.7 ± 8.5	1 week washout of medications	HAM-D(17): 21.9 ± 5.3	No significant differences
**Paralimbic cortical thickness in first-episode depression: evidence for trait-related differences in mood regulation	Van Eijndhoven et al., 2013	MDD (27/13) HC (19/12)	34.95 ± 11.65	Half medication-naïve, during first episode	HAM-D(17): 21.8 ± 4.0 (acutely ill), 3.4 ± 2.0 (recovered)	Thickening in left posterior cingulate cortex (BA 23), left caudal ACC (BA 33), left temporal pole (BA 38) Thinning in left medial OFC (BA 11);

**Local cortical thinning links to resting-state disconnectivity in major depressive disorder	Van Tol et al., 2014	MDD (8/12) HC (3/17)	38.25 ± 11.63	19/20 medicated	HAM-D(17): 15.5 ± 5.9	Thinning in right posterior cingulate cortex, right dorsolateral PFC, right superior temporal gyrus, right dorsomedial PFC
**Prefrontal cortical thickness in depressed patients with high-risk for suicidal behavior	Wagner et al., 2012	MDD-HR (11/4) MDD-nHR (14/1) HC (25/5)	HR: 41.0 ± 12.5 nHR: 34.1 ± 10.5	23/30 medicated (double half-life washout)	HAMD-D(21): 23.9 ± 5.3 (HR), 25.7 ± 5.4 (nHR)	Thinning in left parahippocampal gyrus, left ACC, left OFC, right middle frontal gyrus, right middle temporal gyrus, right superior frontal gyrus, right insula
Regional cortical thinning of the orbitofrontal cortex in medication-naive female patients with major depressive disorder is not associated with MAOA-uVNTR polymorphism	Won et al., 2016	MDD (31/0) HC (43/0)	40.83 ± 9.69	Medication-naïve	HAM-D(17): 20.96 ± 5.09	Thinning in bilateral OFC
**Increased prefrontal and parietal cortical thickness does not correlate with anhedonia in patients with untreated first-episode major depressive disorders	Yang et al., 2015	MDD (13/14) HC (13/14)	28.59 ± 6.82	Medication-naive	HAM-D(24):	Thickening in right medial OFC and left inferior parietal gyrus
**Cortical thickness and subcortical structure volume abnormalities in patients with major depression with and without anxious symptoms.	Zhao et al., 2017a	MDD (20/25) HC (21/22)	32.69 ± 7.85	20/25 medicated or with medication history	HAM-D(17): 24.89 ± 5.51	Thinning in left inferior temporal cortex, right superior temporal cortex, right pars orbitalis
Altered patterns of association between cortical thickness and subcortical volume in patients with first episode major depressive disorder: A structural MRI study	Zhao et al., 2017b	MDD (16/20) HC (18/23)	32.8 ± 8.0	Medication-naïve	HAM-D(17): 25.5 ± 5.3	Thinning in right medial OFC, right inferior temporal cortex, right insula, right inferior parietal region
**Effects of cigarette smoking on cortical thickness in major depressive disorder	Zorlu et al., 2016	MDD-smokers (13/12) MDD-nonsmokers (15/10) HC (14/8)	MDD-smokers: 34.6 ± 10.1 MDD-nonsmokers: 36.7 ± 10.1	Medication-free for 2 months	HAM-D(17): 28.4 ± 4.1 (smokers), 25.9 ± 5.0 (nonsmokers)	Thickening in left middle temporal cortex (extending to superior/inferior gyri), left postcentral gyrus, right insula (in non-smoker MDD compared to HC)

ACC = anterior cingulate cortex; BA = Brodmann area; BDI = Beck Depression Inventory; HAM-D(n) = Hamilton Rating Scale for Depression (number of items); MADRS = Montgomery-Asberg Depression Scale; MDE = major depressive episode; OFC = orbitofrontal cortex; PFC = prefrontal cortex; SI = suicidal ideation; EO = early onset; LO = late onset; HR = high risk; nHR = non-high risk; ns = not specified.
(**) – included in meta-analysis

Table 2. **Summary of neuroimaging parameters and data analysis performed by included studies**

Study	MRI Tesla	Software	Cross-sectional analysis	Multiple comparison correction	Manual edits (Y/N)
Canu et al., 2015	1.5T Philips	FreeSurfer 5.0	Whole-brain, GLM	Monte Carlo	N
Fonseka et al., 2016	3T GE	FreeSurfer (ns)	ROI, SPSS	Not described	N
Fung et al., 2015	3T Siemens	FreeSurfer 5.1.0	Whole-brain, GLM (qdec)	None	Inspection only
Grieve et al., 2013	3T Signa	FreeSurfer 4.3	Whole-brain, Independent sample t-test	FDR	Inspection only
Han et al., 2014	3T Siemens	FreeSurfer 5.0	Whole brain, GLM	Monte Carlo	Inspection only
Han et al., 2017	3T Siemens	FreeSurfer 5.3	Whole brain, Two-way analysis of covariance	FDR	Inspection only
Jarnum et al., 2011	3T Signa	FACE	Whole-brain, One-sided unpaired t-test (unequal var.)	FDR	N
Jaworska et al., 2014	3T Signa	FreeSurfer (ns)	Whole-brain, MANOVA with repeated measures ANOVA	Indirectly by setting uncorrected p<0.01	Y
Lan et al., 2014	3T Signa	FreeSurfer 5.1.0	Whole-brain, GLM (qdec)	Monte Carlo	N
Lener et al., 2016	3T Philips	FreeSurfer (ns)	Whole-brain, two-tailed Student's t-test	Bootstrap	N
Liu et al., 2015	3T Signa	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	Y
Meier et al., 2016	3T GE	FreeSurfer 5.3.0	ROI, MANOVA, MANCOVA	None	Y
Na et al., 2016	3T Siemens	FreeSurfer 5.0	Whole-brain, GLM	Monte Carlo	Inspection only
Niu et al., 2017	3T GE	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	Inspection only

Papmeyer et al., 2015	1.5T Signa	FreeSurfer 5.1.0	ROI, linear mixed-effects model	FDR	N
Peng et al., 2015	3T Signa	FreeSurfer 5.3.0	Whole-brain, GLM	FDR	Y
Perlman et al., 2017	3T various	FreeSurfer 5.3.0	Whole-brain, GLM	Monte Carlo	Inspection only
Phillips et al., 2015	1.5T Siemens	FreeSurfer 4.5	ROI, MANCOVA	None	Y
Pirnia et al., 2016	3T Siemens	FreeSurfer 5.3.0	ROI, GLM	None	Y
Qiu et al., 2014a	3T GE	FreeSurfer 4.5.0	Whole-brain, GLM	FDR	N
Sartorius et al., 2016	3T Siemens	SPM12	Whole-brain, longitudinal only	Family-wise error	N
Spati et al., 2015	3T Philips	FreeSurfer 5.1.0	Whole-brain, GLM	Monte Carlo	N
Taylor et al., 2015	Not specified	FreeSurfer 5.1.0	Whole-brain, GLM	Monte Carlo	Inspection only (edits not needed)
Truong et al., 2013	3T GE	CIVET	Whole-brain, GLM	Random field theory	N
Tu et al., 2012	1.5T GE	FreeSurfer 4.0.5	Whole-brain, GLM	Monte Carlo	N
Van Eijndhoven et al., 2016	1.5T Siemens	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	N
Van Eijndhoven et al., 2013	1.5T Siemens	FreeSurfer (ns)	Whole-brain, GLM (qdec)	None	Y
Van Tol et al., 2014	3T Siemens	CIVET	Whole-brain, two-sample t-test	FDR	N
Wagner et al., 2012	1.5T Siemens	FreeSurfer 5.1.0	Whole-brain, GLM	Monte Carlo	Inspection only
Won et al., 2016	3.0T Siemens	FreeSurfer 5.0	ROI, ANCOVA	Bonferroni	Inspection only
Yang et al., 2015	3T Siemens	FreeSurfer 5.1.0	Whole brain, GLM (qdec)	None	Inspection only
Zhao et al., 2017a	3T Siemens	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	Y
Zhao et al., 2017b	3T Siemens	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	Y
Zorlu et al., 2016	1.5T Philips	FreeSurfer 5.3.0	Whole-brain, GLM (qdec)	Monte Carlo	Y

GLM = general linear model; FDR = false discovery rate; FACE = fast accurate cortical extraction; SPM = Statistical Parametric Mapping; AN(C)OVA = analysis of (co)variance, MAN(C)OVA = multivariate analysis of (co)variance; ns = version not specified.

Table 3. Summary of significant clusters identified by meta-analysis of all included studies (n=24)

Cortical region	Brodmann Area	Size (mm3)	MNI X	MNI Y	MNI Z	SDM-Z	P-value
MDD > HC							
Left supramarginal gyrus	40	140	-46	-44	40	1.229	0.0015
MDD < HC							
Bilateral medial orbitofrontal cortex	11	489	0	30	-28	-1.430	0.00020
Left pars opercularis	45	410	-48	28	0	-1.227	0.00065
Left calcarine fissure/lingual gyrus	17	242	-4	-82	2	-1.328	0.00037

The voxel threshold was p <0.005, peak height threshold was SDM-Z > 1.0 and the extent threshold set at a cluster size of 50 voxels. Regions are labelled according to the FreeSurfer Desikan-Killiany atlas.

Table 4. Summary of significant clusters identified by meta-analysis of subgroup of studies with medication-naïve MDD samples (n=6)

Cortical region	Brodmann Area	Size (mm3)	MNI X	MNI Y	MNI Z	SDM-Z	P-value
MDD > HC							
Right supramarginal gyrus	40	902	56	-48	34	1.531	0.00016
Left inferior parietal gyrus	19	116	-28	-72	34	1.354	0.0018
Right medial orbitofrontal cortex	11	105	6	56	-18	1.246	0.0030
MDD < HC							
Left fusiform gyrus	20	345	-52	-50	-10	-1.208	0.00049
Left lateral orbitofrontal cortex	47	235	-32	22	-10	-1.189	0.00071

Right lateral orbitofrontal cortex	47	168	50	32	-10	-1.208	0.00049
Right middle temporal gyrus	21	58	62	2	-14	-1.191	0.00069

The voxel threshold was p <0.005, peak height threshold was SDM-Z > 1.0 and the extent threshold set at a cluster size of 50 voxels. Regions are labelled according to the FreeSurfer Desikan-Killiany atlas.

Table 5. Egger regression tests and jackknife analyses for each statistically significant cluster from main analyses

Cortical region	Brodmann Area	Egger's bias indicator	T-statistic	P-value	Jackknife analysis
Overall analysis (df = 22)					
MDD > HC					
Left supramarginal gyrus	40	-1.35	-1.65	0.113	23/24
MDD < HC					
Bilateral medial orbitofrontal cortex	11	0.23	0.48	0.636	23/24
Left pars opercularis	45	0.23	0.39	0.701	23/24
Left calcarine fissure/lingual gyrus	17	0.62	0.91	0.371	23/24
Subgroup analysis (df = 4)					
MDD > HC					
Right supramarginal gyrus	40	-0.76	-0.30	0.777	1/6
Left inferior parietal gyrus	19	3.67	1.55	0.197	1/6
Right medial orbitofrontal cortex	11	-2.33	-1.02	0.367	3/6
MDD < HC					
Left fusiform gyrus	20	1.98	0.81	0.463	5/6
Left lateral orbitofrontal cortex	47	-0.12	-0.05	0.965	5/6
Right lateral orbitofrontal cortex	47	1.99	0.79	0.471	3/6
Right middle temporal gyrus	21	2.19	0.96	0.391	1/6

The ratio for jackknife analyses indicates the number of studies in which the finding was replicated, out of all total studies included in the analysis. df = degrees of freedom.

Table 6. Summary of correlations identified between cortical thickness and clinical variables in the MDD samples

Summary of correlations found between cortical thickness in MDD and clinical variables

General clinical severity	HAM-D score (+) **L supramarginal gyrus** thickness, (-) left lingual cortex thickness (Na et al., 2016)	BDI score (+) **bilateral superior parietal gyri** thickness (Yang et al., 2015)	Duration of illness (-) **R medial OFC** thickness (Yang et al., 2015)	HAM-D score (-) **R rostral middle frontal gyrus** and **supramarginal gyrus** (increased in MDD) (Qiu et al., 2014)	HAM-D score (+) **posterior cingulate cortex** thickness (Truong et al., 2013)	HAM-D score (-) thickness of **L precentral gyrus** and **L superior temporal gyrus** (Zorlu et al., 2016)
Symptom dimensions	HAM-A respiratory subscore (-) **R medial OFC** thickness (Canu et al., 2016)	Visual analog scores for irritation and fatigue (-) **R rostral ACC** thickness (Lener et al., 2015)	GAD comorbidity (+) greater thinning in **frontotemporal cortex** (Canu et al., 2015)	Time required to complete CTT2 (worse executive performance) (-) thickness of **R rostral middle frontal gyrus**, **R superior frontal gyrus** and **R supramarginal gyrus** (Tu et al., 2011)	Trait anxiety score (-) **L medial OFC** thickness (van Eijndhoven et al., 2013)	
Vulnerability and risk	CTQ (-) left precuneus and bilateral frontal pole thickness; abuse subscore (-) **R frontal pole** thickness and neglect subscore (-) **L inferior and right superior parietal gyri** (Jaworska et al., 2014)	Family history of mood disorder (+) **OFC** thickness (Lan et al., 2015)				

Physiological measures/genotype	FKBP5 methylation (+) **R transverse frontopolar gyrus** thickness in C allele homozygote group (Han et al., 2017)	Serum cortisol (-) thickness of **lateral and midline frontal regions** and **bilateral temporal regions** (Liu et al., 2015)	BDNF promoter methylation at CpG2 (higher in MDD sample) and CpG4 (-) thickness in several regions encompassing **all four lobes** (Na et al., 2016)		
Clinical response (longitudinal)	Lower HAM-D score (+) thickness of **R insula** (van Eijndhoven et al., 2016)	Lower MADRS score following antidepressant treatment (+) thickness of **R caudal ACC** at baseline (Phillips et al., 2015)	Overall clinical response following ECT (+) **ACC** thickness at baseline (Pirnia et al., 2016)	Number of ECT sessions (+) post-treatment thickness in three regions of **L temporal cortex** (van Eijndhoven et al., 2016)	Responders to ECT showed larger increase in thickness compared to non-responders (van Eijndhoven, 2016)

(+) indicates a positive correlation and (-) indicates a negative correlation. ACC = anterior cingulate cortex; BDI = Beck Depression Inventory; BDNF = brain-derived neurotrophic factor; CTQ = Childhood Trauma Questionnaire; CTT = Color Trails Test; ECT = electroconvulsive therapy; FKBP5 = FK506 binding protein 51; GAD = generalized anxiety disorder; HAM-A = Hamilton Rating Scale for Anxiety; HAM-D = Hamilton Rating Scale for Depression; MADRS = Montgomery-Asberg Depression Scale; OFC = orbitofrontal cortex.

Table 7. A comparison of the significant clusters identified in the ENGIMA study and the current meta-analysis

Areas identified in ENIGMA study	Areas identified in current study	
MDD < HC	*MDD < HC*	*MDD > HC*
Overall group analysis		
Bilateral medial OFC**	Bilateral medial OFC**	Left supramarginal gyrus
Bilateral fusiform gyri	Left pars opercularis	
Bilateral insula	Left calcarine fissure/lingual gyrus	
Bilateral rostral ACC		
Bilaterial posterior cingulate cortex		
Left middle temporal gyrus		
Right inferior temporal gyrus		
Right caudal ACC		
First episode subgroup analysis		
Bilaterial fusiform gyri**	Left fusiform gyrus**	Right supramarginal gyrus
Bilateral rostral ACC	Bilateral OFC	Left inferior parietal gyrus
Bilateral insula	Right middle temporal gyrus	Right medial OFC
Left medial OFC		
Left superior frontal cortex		
Right caudal ACC		
Right caudal posterior cingulate cortex		
Right isthmus cingulate cortex		

Areas that have been identified by both studies have been indicated by double asterisks (**). Results with the same label but in opposing hemispheres were not considered common to both studies. All labels derive from the FreeSurfer Desikan-Killiany atlas.

Chapter 3: Cerebral cortical thickness after treatment with desvenlafaxine succinate in major depressive disorder

Jee Su Suh [1,2]; Luciano Minuzzi [1,2,3]; Lauren E. Cudney [2,4], William Maich [2], Maha Eltayebani [2,5], Claudio N. Soares [6,7], Benicio N. Frey [1,2,3]

[1] Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada
[2] Mood Disorders Program and Women's Health Concerns Clinic, St. Joseph's Healthcare Hamilton, ON, Canada
[3] Department of Psychiatry, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil
[4] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada
[5] Department of Neuropsychiatry, Faculty of Medicine, Alexandria University, Alexandria, Egypt
[6] Department of Psychiatry, Queen's University School of Medicine, Kingston, ON, Canada
[7] St. Michael's Hospital, University of Toronto, Toronto, ON, Canada

The chapter in its entirety has been **published** in **NeuroReport**. The final accepted manuscript version of this article is presented within this thesis.

Abstract

Thickness of the cerebral cortex has been previously investigated for its potential as a biomarker in major depressive disorder (MDD). This is the first study to examine the longitudinal effects of a serotonin-norepinephrine reuptake inhibitor, desvenlafaxine succinate (DVS), on whole-brain cortical thickness in patients treated for MDD. We also aimed to replicate a previous finding of an association between improvement in clinical severity and cortical thickness in five predefined regions-of-interest (ROI). Twenty-five individuals with MDD received treatment with DVS (50 mg/d) for 8 weeks, with 19 completing the study. We used FreeSurfer 6.0 to compare group differences between MDD and controls (n=23) and between treatment responders, treatment non-responders and controls. We tested correlations between 8-week change in depression severity and regional cortical thickness in five ROIs: the rostral and caudal anterior cingulate cortices, lateral and medial orbitofrontal cortices and inferior temporal gyrus. There were no differences in cortical thickness between MDD and controls or DVS responders and controls. There was cortical thickening in DVS non-responders in the left pars orbitalis when compared to controls (MNI[X,Y,Z] = [-38.4, 37.6, -11.1]; p-value = 0.027). There were no significant correlations between change in depression severity and cortical thickness in any of the five ROIs. Brain cortical thickness does not seem to be a sensitive marker of short-term antidepressant response in MDD, except increased cortical thickness in non-responders. Duration of the intervention and inter-individual heterogeneity may impede identification of discriminating features of treatment response as associated to cortical thickness.

3.1 Introduction

Major depressive disorder (MDD) involves changes in mood, cognition, sleep and psychomotor function, which can vary greatly between individuals. Given limited remission rates after antidepressant treatment [1], there is an ongoing search to identify and describe biomarkers that could be used to define subtypes among a heterogenous pool of patients [2–4]. The goal is to define an optimal combination of biomarkers that can reliably and meaningfully inform which individuals are more likely to respond to antidepressant treatment.

As a consequence of this heterogeneity, the literature on the neuroimaging correlates of MDD is extensive and varied. Functional studies have identified widespread network dysfunction, particularly in the default mode and cognitive control networks comprising all four cerebral lobes [5]. Given these diffuse effects, abnormalities in cortical thickness (CT) are difficult to elucidate [6]. Thinner cortex has been seen in both unmedicated MDD and high-risk samples, which could indicate vulnerability to the disorder [7]. However, results are varied, ranging from thinning in all identified regions to both thinner and thicker cortex in the MDD group as well as null results. Given this lack of consensus, there is a need for additional data supported by rigorous statistical procedures.

It has been suggested that pharmacological antidepressant treatment may 'normalize' thinned cortex in MDD, correlating with symptom improvement [8,9]. However, animal studies examining possible causal links between first-line treatment and CT have observed a lack of change in thickness following drug administration [10,11].

Moreover, this hypothesis has never been tested in a study using the antidepressant desvenlafaxine succinate (DVS), a serotonin-noradrenaline reuptake inhibitor (SNRI).

Furthermore, associations between baseline CT and treatment response at varying time-points are emerging [8,9,12], although results are mixed. Other longitudinal studies of CT have examined the effects of various treatments including a selective serotonin reuptake inhibitor (SSRI) as well as nonspecific pharmacotherapy [8,9,13–16]. One study found that among five predefined regions-of-interest (ROI), thicker caudal anterior cingulate cortex (ACC) was correlated to symptom improvement over the treatment course involving varied pharmacotherapies [9]. There are no studies to date examining whether baseline CT can predict response to a specific SNRI, a gap we are aiming to address using a DVS treatment study with a sample of mid-life men and women.

We used two analytic approaches to address both exploratory and confirmatory aspects. Our primary objectives were to determine, using an exploratory whole-brain model-free approach, (a) whether certain CT features at baseline are associated to clinical response following 8 weeks of DVS treatment, (b) whether there are any structural changes within the MDD group as a result of the use of this antidepressant with or without satisfactory treatment response and (c) to compare baseline CT between MDD and healthy controls (HCs). A secondary objective was to use an a priori region- based approach, to replicate previous findings in Phillips et al. [9] described above. We will use the same five ROI to test correlations between baseline thickness in these regions and change in depression severity over the treatment period.

3.2 Methods

3.2.1 Participants

Participants were recruited through the Mood Disorders Program and the Women's Health Concerns Clinic at St Joseph's Healthcare, Hamilton, Ontario. All participants provided signed written informed consent before study entry. This study was approved by the local Research Ethics Board.

Inclusion criteria for MDD participants included: (a) 40–60 years of age; (b) diagnosis of current MDD according to the Mini-International Neuropsychiatric Interview (MINI); (c) score of at least 25 on the Montgomery–Åsberg Depression Rating Scale (MADRS). Exclusion criteria were: (a) current axis I psychiatric disorder other than MDD assessed by the MINI; (b) regular treatment with any SSRI or SNRI or other agents known to influence mood within 4 weeks before screening visit; (c) suicidal or homicidal ideation, psychotic symptoms; (d) laboratory abnormalities; (e) presence of significant medical issues; (f) pregnancy or breastfeeding. Exclusion criteria for age-matched and sex-matched HCs were as follows: (a) any lifetime axis I diagnosis according to the MINI; (b) substance abuse in the past 6 months; (c) presence of significant neurological disorder; (d) presence of MRI contraindications; (e) pregnant/breastfeeding.

3.2.2 Experimental design

Treatment consisted of a 2-week placebo lead-in, followed by an 8-week open-label trial with DVS. After the placebo lead-in phase, placebo responders (defined as $\geq$ 50% decrease in MADRS score), were excluded. Placebo non- responders started an open

trial with DVS (50 mg/day). At each visit (baseline, weeks 2, 4, 6, and 8), mood was assessed with clinician-rated MADRS. Self-report scales were completed at baseline and week 8 by all participants. Sleep quality was assessed with the Pittsburgh Sleep Quality Index, quality of life assessed using the abbreviated WHO Quality of Life questionnaire (BREF), anxiety symptom severity measured with the Beck Anxiety Index and cognitive errors measured with the Cognitive Failures Questionnaire.

3.2.3 Brain imaging and statistical analyses

Clinical measures were compared between MDD and HCs and changes from baseline to week 8 were considered using a within-subject analysis. Paired t-tests were used for normally distributed data. Wilcoxon matched pairs signed rank and Mann–Whitney U-tests were used for data that was not normally distributed.

Brain imaging was performed at the baseline and week 8 visits using a GE 3T whole body short bore scanner with parallel receiver channels (General Electric, Milwaukee, Wisconsin, USA). High-resolution T1-weighted anatomical images were acquired using a gradient-echo inversion-recovery sequence (TR = 1.6 s, TE = 5 ms, matrix 256×256×128, FOV 220×220mm, slice thick- ness 1 mm). T1 images were processed for CT analysis using FreeSurfer 6.0 (https://surfer.nmr.mgh.harvard.edu/) using a fully automated procedure [17]. The processed images were inspected visually for segmentation errors; a researcher blinded to group assignments carried out minor manual edits. The final images were smoothed with a 10 mm full-width half-maximum Gaussian kernel.

To address our primary exploratory objectives, general linear modelling from the FreeSurfer suite was used to model whole brain, vertex-wise differences in CT between MDD subgroups (pooled, responder, and nonresponder) and HC as well as longitudinal changes within the MDD group, controlled for age. Whole-brain results were corrected for multiple comparisons using cluster-wise thresholding with a vertex-wise threshold of P=0.01 and a cluster-forming threshold of P=0.05 [18], with an additional Bonferroni correction for two-hemisphere testing.

For our secondary confirmatory analyses, Python 3.6.3 was used to test correlations between 8-week change in MADRS score and baseline average CT values for each of the following ROIs as defined by the FreeSurfer Desikan–Killiany atlas, in both hemispheres: rostral and caudal ACC, lateral and medial orbitofrontal cortex and inferior temporal gyrus. The atlas and regions are the same as those investigated by the authors of the study whose results we are aiming to replicate [9]. Pearson's correlations were calculated for normally distributed variables while Spearman correlations were used for non-normal variables and were corrected using the false discovery rate at P = 0.05.

3.3 Results

Following the placebo lead-in phase, 25 eligible participants started treatment with DVS. Six participants were excluded because of attrition or incomplete imaging. Data passed quality checks for 19 MDD participants and 23 HC at baseline. We obtained week 8 response status for 16 patients, of which 12 were DVS responders. Thirteen MDD

participants had neuroimaging data at week 8 for within-group longitudinal analyses, of which nine were DVS responders.

3.3.1 Clinical outcomes measures

Treatment with DVS led to significant improvements in depressive symptoms (MADRS), sleep quality (Pittsburgh Sleep Quality Index) and quality of life (WHO Quality of Life) (Table 1). Cognitive Failures Questionnaire and Beck Anxiety Index did not significantly change following treatment (both $P > 0.05$).

3.3.2 Cortical thickness

Whole-brain analyses revealed that the left pars orbitalis, also known as the orbital part of the inferior frontal gyrus (BA47), was found to be thicker in the nonresponder group when compared to HC (Fig. 1 and Table 2). No differences in baseline CT were observed between pooled MDD versus HC or responders versus HC. No significant longitudinal changes were seen within any of the above groups. There were no significant correlations between CT in any of the ROIs and change in MADRS score.

3.4 Discussion

Our data revealed CT features that can distinguish MDD nonresponders from controls at baseline, exhibited as thickening in a region relevant to the circuitry of the disorder. This result is difficult to validate on the basis of existing literature because of the lack of studies containing equivalent nonresponder versus HC comparisons.

Moreover, tentative explanations for results of thicker cortex in MDD have not been confirmed, although it has been suggested that thickening may be a marker of inflammation due to compensatory mechanisms in response to the onset of a major depressive episode [6]. The pars orbitalis has been implicated primarily in emotional regulation as well as in executive function [19]. However, it is unlikely that thickening in this region is specific to MDD; other studies have also identified thickening in the orbitofrontal region in other neuropsychiatric disorders such as social anxiety and autism spectrum disorders [20,21], indicative of its larger role in emotional and executive dysfunction. Moreover, given the small sample size from which we derived our result, we cannot reach a definite conclusion as to why this specific result was seen. It is possible that greater pre- treatment structural deviation from HC may suggest the need for a longer treatment time to respond to DVS or a preferential response to a different and/or adjunctive treatment. A recent multisite study did not find any significant associations between baseline CT and treatment response to a SSRI [8]. However, their ROI analysis was restricted to a few regions that did not include the pars orbitalis, and therefore could not be used to corroborate or refute our results.

The whole-brain comparison between the pooled MDD and HC groups revealed no differences in CT at baseline, despite previous positive findings [6]. This variation of results is likely due to heterogeneity in this diagnostic population. Larger sample sizes and subtyping are required to further reduce the heterogeneity present in the overall sample. The recent publication by the ENIGMA consortium (N > 10 000) revealed only small effect sizes for CT differences between adult MDD and HC [22]. With our current

sample size, we were not able to address this particular limitation, but instead utilized an exploratory, model-free approach with a multi-step protocol for multiple corrections in order to produce statistically reliable results. Additionally, we found no significant longitudinal changes in whole-brain CT during the treatment course. It may be that the clinical changes led by DVS are not reflected by changes in CT or that an assessment after only 8 weeks of treatment is too soon to detect CT changes induced by DVS.

Given the large variation of CT results in the literature for MDD, we attempted to reproduce results from a highly-cited publication with a comparable sample size [9]. We were not able to replicate their positive result (thicker baseline caudal ACC associated to symptom improvement), although this could be due to several reasons. First, antidepressant use was not controlled for in the previous study, whereas our current study was a case-controlled trial with standardized administration of DVS. Second, differences in mean ages and symptom severity may have contributed to the discrepancy between the two sets of results. Additionally, sample sizes in both studies were small, with a total N ranging from 42 to 54 (including HC), increasing the chances of spurious findings due to reduced power [23].

This paper is the first in the literature to examine cross- sectional and longitudinal CT features in the context of a specific SNRI, which is the main strength of the study. Automatically processed images were manually edited following blinding to correct for minor segmentation errors, improving the overall quality of the dataset [24]. Our limitations include the small sample size, particularly in the nonresponder group (n=4); in the field of neuroscience in particular, studies with small sample sizes were found to

overestimate the true effect [23]. We attempted to mitigate this limitation by using rigorous statistical thresholds for all analyses.

3.5 Conclusion

We have shown here, with a sample comparable to others found in the literature, that SNRI nonresponders had increased thickness in the orbital part of the inferior frontal gyrus (BA47). Further research investigating whole-brain differences between these groups is needed, as ROI-based approaches commonly exclude regions outside of the ACC and orbitofrontal cortex. The lack of replicability of correlations between specific ROIs and clinical improvement and the lack of differences between HC and the overall MDD group suggest variance among patients that may conceal group differences of smaller effect size. Using larger sample sizes, rigorous statistical controls for whole-brain analyses and subtyping of MDD groups are needed to counteract these limitations. Recent unsupervised machine learning approaches have shown the feasibility of subtyping on the basis of neuroimaging measures [25], although these subtypes have yet to be replicated or confirmed from a clinical perspective. Future studies might attempt to combine multivariate and univariate analyses to extract the complex patterns of clinical symptomology and objective biomarkers that exist under the heterogeneous umbrella label of MDD.

3.6 Funding

3.7 Declaration of Interest

Benicio Frey has served on advisory boards for Lundbeck and Sunovion and received research grant from Pfizer Canada. Claudio Soares has served on advisory boards for Lundbeck, Sunovion, Pfizer, Servier, Merck and Otsuka. The remaining authors have no conflicts of interest to declare.

3.8 References

Bartlett, E. A., DeLorenzo, C., Sharma, P., Yang, J., Zhang, M., Petkova, E., Weissman, M., McGrath, P. J., Fava, M., Ogden, R. T., Kurian, B. T., Malchow, A., Cooper, C. M., Trombello, J. M., McInnis, M., Adams, P., Oquendo, M. A., Pizzagalli, D. A., Trivedi, M., & Parsey, R. V. (2018). Pretreatment and early-treatment cortical thickness is associated with SSRI treatment response in major depressive disorder. Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 43(11), 2221–2230. https://doi.org/10.1038/s41386-018-0122-9

Belyk, M., Brown, S., Lim, J., & Kotz, S. A. (2017). Convergence of semantics and emotional expression within the IFG pars orbitalis. NeuroImage, 156, 240–248. https://doi.org/10.1016/j.neuroimage.2017.04.020

Boes, A. D., Uitermarkt, B. D., Albazron, F. M., Lan, M. J., Liston, C., Pascual-Leone, A., Dubin, M. J., & Fox, M. D. (2018). Rostral anterior cingulate cortex is a structural correlate of repetitive TMS treatment response in depression. Brain Stimulation, 11(3), 575–581. https://doi.org/10.1016/j.brs.2018.01.029

Brühl, A. B., Hänggi, J., Baur, V., Rufer, M., Delsignore, A., Weidt, S., Jäncke, L., & Herwig, U. (2014). Increased cortical thickness in a frontoparietal network in social anxiety disorder. Human Brain Mapping, 35(7), 2966–2977. https://doi.org/10.1002/hbm.22378

Brunoni, A. R., Sampaio-Junior, B., Moffa, A. H., Borrione, L., Nogueira, B. S., Aparicio, L. V. M., Veronezi, B., Moreno, M., Fernandes, R. A., Tavares, D.,

Bueno, P. V. S., Seibt, O., Bikson, M., Fraguas, R., & Bensenor, I. M. (2015). The Escitalopram versus Electric Current Therapy for Treating Depression Clinical Study (ELECT-TDCS): Rationale and study design of a non-inferiority, triple-arm, placebo-controlled clinical trial. Sao Paulo Medical Journal, 133(3), 252–263. https://doi.org/10.1590/1516-3180.2014.00351712

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience, 14(5), 365–376. https://doi.org/10.1038/nrn3475

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R. N., Zebley, B., Oathes, D. J., Etkin, A., Schatzberg, A. F., Sudheimer, K., Keller, J., Mayberg, H. S., Gunning, F. M., Alexopoulos, G. S., Fox, M. D., Pascual-Leone, A., Voss, H. U., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nature Medicine, 23(1), 28–38. https://doi.org/10.1038/nm.4246

Fischl, B. (2012). FreeSurfer. NeuroImage, 62(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fu, C. H. Y., Costafreda, S. G., Sankar, A., Adams, T. M., Rasenick, M. M., Liu, P., Donati, R., Maglanoc, L. A., Horton, P., & Marangell, L. B. (2015). Multimodal functional and structural neuroimaging investigation of major depressive disorder following treatment with duloxetine. BMC Psychiatry, 15, 82. https://doi.org/10.1186/s12888-015-0457-2

Hagler, D. J., Saygin, A. P., & Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. NeuroImage, 33(4), 1093–1103. https://doi.org/10.1016/j.neuroimage.2006.07.036

Hao, X., Talati, A., Shankman, S. A., Liu, J., Kaiser, J., Tenke, C. E., Warner, V., Semanek, D., Wickramaratne, P. J., Weissman, M. M., & Posner, J. (2017). Stability of Cortical Thinning in Persons at Increased Familial Risk for Major Depressive Disorder Across 8 Years. Biological Psychiatry. Cognitive Neuroscience and Neuroimaging, 2(7), 619–625. https://doi.org/10.1016/j.bpsc.2017.04.009

Lam, R. W., Milev, R., Rotzinger, S., Andreazza, A. C., Blier, P., Brenner, C., Daskalakis, Z. J., Dharsee, M., Downar, J., Evans, K. R., Farzan, F., Foster, J. A., Frey, B. N., Geraci, J., Giacobbe, P., Feilotter, H. E., Hall, G. B., Harkness, K. L., Hassel, S., … Team, C.-B. I. (2016). Discovering biomarkers for antidepressant response: Protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. Bmc Psychiatry, 16, 105. https://doi.org/10.1186/s12888-016-0785-x

Lyttle, K., Ohmura, Y., Konno, K., Yoshida, T., Izumi, T., Watanabe, M., & Yoshioka, M. (2015). Repeated fluvoxamine treatment recovers juvenile stress-induced morphological changes and depressive-like behavior in rats. Brain Research, 1616, 88–100. https://doi.org/10.1016/j.brainres.2015.04.058

Phillips, J. L., Batten, L. A., Tremblay, P., Aldosary, F., & Blier, P. (2015). A

Prospective, Longitudinal Study of the Effect of Remission on Cortical

Thickness and Hippocampal Volume in Patients with Treatment-Resistant

Depression. International Journal of Neuropsychopharmacology, 18(8).

https://doi.org/10.1093/ijnp/pyv037

Phillips, M. L., Chase, H. W., Sheline, Y. I., Etkin, A., Almeida, J. R. C., Deckersbach,

T., & Trivedi, M. H. (2015). Identifying Predictors, Moderators, and Mediators

of Antidepressant Response in Major Depressive Disorder: Neuroimaging

Approaches. The American Journal of Psychiatry, 172(2), 124–138.

https://doi.org/10.1176/appi.ajp.2014.14010076

Pirnia, T., Joshi, S. H., Leaver, A. M., Vasavada, M., Njau, S., Woods, R. P., Espinoza,

R., & Narr, K. L. (2016). Electroconvulsive therapy and structural

neuroplasticity in neocortical, limbic and paralimbic cortex. Translational

Psychiatry, 6(6), e832. https://doi.org/10.1038/tp.2016.102

Popescu, V., Klaver, R., Versteeg, A., Voorn, P., Twisk, J. W. R., Barkhof, F., Geurts, J.

J. G., & Vrenken, H. (2016). Postmortem Validation of MRI Cortical Volume

Measurements in MS. Human Brain Mapping, 37(6), 2223–2233.

https://doi.org/10.1002/hbm.23168

Sartorius, A., Demirakca, T., Böhringer, A., Clemm von Hohenberg, C., Aksay, S. S.,

Bumb, J. M., Kranaster, L., & Ende, G. (2016). Electroconvulsive therapy

increases temporal gray matter volume and cortical thickness. European

Neuropsychopharmacology: The Journal of the European College of

Neuropsychopharmacology, 26(3), 506–517.

https://doi.org/10.1016/j.euroneuro.2015.12.036

Schmaal, L., Hibar, D. P., Saemann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N.,

Cheung, J. W., Erp, T. G. M. van, Bos, D., Ikram, M. A., Vernooij, M. W.,

Niessen, W. J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H. J., Janowitz,

D., Buelow, R., Selonke, M., … Depressive, E. M. (2017). Cortical

abnormalities in adults and adolescents with major depression based on brain

scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder

Working Group. Molecular Psychiatry, 22(6), 900–909.

https://doi.org/10.1038/mp.2016.60

Suh, J. S., Schneider, M. A., Minuzzi, L., MacQueen, G. M., Strother, S. C., Kennedy, S.

H., & Frey, B. N. (2018). Cortical thickness in major depressive disorder: A

systematic review and meta-analysis. Progress in Neuro-Psychopharmacology &

Biological Psychiatry, 88, 287–302. https://doi.org/10.1016/j.pnpbp.2018.08.008

Trivedi, M. H., McGrath, P. J., Fava, M., Parsey, R. V., Kurian, B. T., Phillips, M. L.,

Oquendo, M. A., Bruder, G., Pizzagalli, D., Toups, M., Cooper, C., Adams, P.,

Weyandt, S., Morris, D. W., Grannemann, B. D., Ogden, R. T., Buckner, R.,

McInnis, M., Kraemer, H. C., … Weissman, M. M. (2016). Establishing

moderators and biosignatures of antidepressant response in clinical care

(EMBARC): Rationale and design. Journal of Psychiatric Research, 78, 11–23.

https://doi.org/10.1016/j.jpsychires.2016.03.001

van Eijndhoven, P., Mulders, P., Kwekkeboom, L., van Oostrom, I., van Beek, M., Janzing, J., Schene, A., & Tendolkar, I. (2016). Bilateral ECT induces bilateral increases in regional cortical thickness. Translational Psychiatry, 6(8), e874. https://doi.org/10.1038/tp.2016.139

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., Calderoni, S., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F. L. S., Durston, S., Ecker, C., Fair, D., Fedor, J., Fitzgerald, J., Freitag, C. M., Gallagher, L., … Buitelaar, J. K. (2017). Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. American Journal of Psychiatry, 175(4), 359–369. https://doi.org/10.1176/appi.ajp.2017.17010100

Wade, B. S. C., Sui, J., Hellemann, G., Leaver, A. M., Espinoza, R. T., Woods, R. P., Abbott, C. C., Joshi, S. H., & Narr, K. L. (2017). Inter and intra-hemispheric structural imaging markers predict depression relapse after electroconvulsive therapy: A multisite study. Translational Psychiatry, 7(12), 1270. https://doi.org/10.1038/s41398-017-0020-7

Williams, L. M. (2017). Getting Personalized: Brain Scan Biomarkers for Guiding Depression Interventions. American Journal of Psychiatry, 174(6), 503–505. https://doi.org/10.1176/appi.ajp.2017.17030314

3.9 Figures and Tables

Figure 1. The left pars orbitalis (BA 47) was found to be thicker in MDD non-responders to DVS as compared to HC in a whole-brain analysis. The scale refers to the cluster-wise -log(p) value.

Table 1. **Characteristics of study participants.**

Measures	MDD (n=19)		HC (n=23)
	Baseline	Week 8	
Age, *Mean (SD)*	53.8 (3.8)	--	50.7 (5.6)†
Sex, *N (%)*			
Male	15 (78.9)	--	13 (56.5)
Female	4 (21.1)	--	10 (43.5)
Handedness, *N (%)*			
Right	16 (84.2)	--	22 (95.6)
Left	0 (0)	--	1 (4.3)
Ambidextrous	3 (15.7)	--	0 (0)
MADRS, *Mean (SD)*	25.0 (6.6)	11.7 (9.1)**	1.6 (2.3)††
Beck Anxiety Index, *Mean (SD)*	13.5 (6.6)	11.4 (10.8)	2.9 (3.1)††
Pittsburgh Sleep Quality Index, *Mean (SD)*	11.8 (3.7)	10.0 (4.7)*	4.9 (3.9)††
Cognitive Failures Questionnaire, *Mean (SD)*	47.5 (17.5)	43.9 (20.0)	24.0 (13.6)††
WHO Quality of Life, *Mean (SD)*	159.1 (62.7)	185.9 (51.0)*	288.7 (50.6)††

Asterisks (*) indicate significance of the within-group comparison from baseline to week 8 in the MDD group. Crosses (†) indicate significance of the between-group comparison at baseline between MDD and HC. */† - $p < 0.05$; **/†† - $p < 0.001$

Table 2. **Cluster information for comparison of whole-brain cortical thickness between MDD non-responders and HC.**

ClusterNo	Max Value	Annotation	Size(mm^2)	MNI(X)	MNI(Y)	MNI(Z)	CWP
1	4.663	Left pars orbitalis	429.33	-38.4	37.6	-11.1	0.027

Chapter 4: An investigation of cortical thickness and antidepressant response in major depressive disorder: a CAN-BIND study report

Jee Su Suh, BSc[a,b,] ; Luciano Minuzzi, MD PhD[a,b,c]; Pradeep Reddy Raamana, PhD[d], Andrew Davis, PhD[e], Geoffrey B. Hall, PhD[e], Jacqueline Harris, MSc[f], Stefanie Hassel, PhD[g,h], Mojdeh Zamyadi, MSc[d], Stephen R. Arnott PhD[d], Gésine L. Alders, PhD[a,b], Roberto B. Sassi, MD PhD[c], Roumen Milev MD PhD[i], Raymond W. Lam, MD[j], Glenda M. MacQueen, MD PhD[g,h], Stephen C. Strother, PhD[d], Sidney H. Kennedy, MD[k,l], Benicio N. Frey, MD PhD[a,b,c]

[a] Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada
[b] Mood Disorders Program and Women's Health Concerns Clinic, St. Joseph's Healthcare Hamilton, ON, Canada
[c] Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada
[d] Rotman Research Institute, Baycrest Health Sciences; Department of Medical Biophysics, University of Toronto, ON, Canada
[e] Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada
[f] Department of Computing Science, University of Alberta, Edmonton, AB, Canada
[g] Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[h] Mathison Centre for Mental Health Research and Education, University of Calgary, Calgary, AB, Canada
[i] Departments of Psychiatry and Psychology, Queen's University and Providence Care Hospital, Kingston, ON, Canada
[j] Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada
[k] Canadian Biomarker Integration Network for Depression, St. Michael's Hospital, Toronto, ON, Canada
[l] Department of Psychiatry, University of Toronto, Toronto, ON, Canada

The chapter in its entirety has been **published** in **NeuroImage: Clinical**. The final accepted manuscript version of this article is presented within this thesis.

Abstract

Major depressive disorder (MDD) is considered a highly heterogeneous clinical and neurobiological mental disorder. We employed a novel layered treatment design to investigate whether cortical thickness features at baseline differentiated treatment responders from non-responders after 8 and 16 weeks of a standardized sequential antidepressant treatment. Secondary analyses examined baseline differences between MDD and controls as a replication analysis and longitudinal changes in thickness after 8 weeks of escitalopram treatment. 181 MDD and 95 healthy comparison (HC) participants were studied. After 8 weeks of escitalopram treatment (10-20mg/d, flexible dosage), responders (>50% decrease in Montgomery-Åsberg Depression Scale score) were continued on escitalopram; non-responders received adjunctive aripiprazole (2-10mg/d, flexible dosage). MDD participants were classified into subgroups according to their response profiles at weeks 8 and 16. Baseline group differences in cortical thickness were analyzed with FreeSurfer between HC and MDD subgroups as well as between response groups. Two-stage longitudinal processing was used to investigate 8-week escitalopram treatment-related changes in cortical thickness. Compared to HC, the MDD group exhibited thinner cortex in the left rostral middle frontal cortex [MNI(X,Y,Z=−29,9,54.5,−7.7); CWP=0.0002]. No baseline differences in cortical thickness were observed between responders and non-responders based on week-8 or week-16 response profile. No changes in cortical thickness were observed after 8 weeks of escitalopram monotherapy. In a two-step 16-week sequential clinical trial we found

that baseline cortical thickness does not appear to be associated to clinical response to pharmacotherapy at 8 or 16 weeks.

4.1 Introduction

Major depressive disorder (MDD) affects up to 300 million people worldwide and is one of the most prevalent causes of disability globally ("WHO | Depression" n.d.). First-line antidepressants have limited efficacy (Cipriani et al. 2018), often necessitating additional treatment courses that can prolong or worsen the patient's distress. Moreover, antidepressants are prescribed based on group average responses from clinical trials, rather than on objective individual characteristics derived from clinical or neurobiological data. The high degree of heterogeneity among individuals meeting diagnostic criteria for MDD likely underlie a wide range of neurobiological subtypes (Drysdale et al. 2017). Clinical- and biomarker-informed treatment selection is the ultimate goal of precision medicine, where accurate subtyping would help clinicians discern whether certain medications are differentially effective in a subtype-dependent manner (Paris 2014; Young et al. 2016; Kennedy and Ceniti 2018). The Canadian Biomarker Integration Network in Depression (CAN-BIND) is a multi-site clinical treatment trial involving several major research centres in Canada (Kennedy and Ceniti 2018; Lam et al. 2016). Clinical, molecular and neuroimaging data were collected from over 300 participants, including MDD patients and healthy comparison (HC) participants.

Neuroimaging has emerged as a promising approach in the search for biomarkers, including anatomical magnetic resonance imaging (MRI), which makes possible the visualization and quantification of the structure of the brain at millimeter resolution. One such parameter is cortical thickness, defined as the distance from the pial boundary to the gray matter (GM)/white matter (WM) boundary comprising the cell bodies of the cerebral

cortex as well as intracortical myelin (Narr et al. 2007; Seldon 2007; Shaw et al. 2008).

There are various factors that could lead to cortical thinning, including reduction of

synapses, atrophy of dendritic trees or reduced vascularization (Lyttle et al. 2015; Schüz

and Palm 1989). Cortical thinning, independent of aging processes, has been

demonstrated in samples of MDD participants and has been replicated in previous studies,

albeit not without ambiguity regarding the specific subregions affected by cortical

thinning (Schmaal et al. 2017; Pink et al. 2017; Suh, Schneider, et al. 2019). Moreover,

cortical thickening has also been observed, particularly in medication-naïve, first-episode

MDD patients (Philip van Eijndhoven et al. 2013; Fonseka et al. 2016; Peng et al. 2015;

Qiu et al. 2014; Yang et al. 2015), which has been hypothesized to reflect glial

hypertrophy as an immune response to initial excitotoxic injury during the first episode

(Dowlati et al. 2010).

　　　Until recently (Schmaal et al. 2017; Perlman et al. 2017), studies on cortical

thickness were hampered by small-to-moderate sample sizes (N<100) and therefore low

power, especially when the statistical considerations that must be made for multi-

dimensional neuroimaging data are taken into account (Cremers, Wager, and Yarkoni

2017). Under-powered neuroimaging studies suffer from effect size inflation and have

low replicability (Button et al. 2013), often constraining statistical analyses to regions of

interest that may not fully capture the whole-brain signature of the associated disorder.

Another limitation is the paucity of longitudinal studies assessing patient response to a

given antidepressant over time (Suh et al. 2019b). The few longitudinal studies that have

tracked changes in cortical thickness over time are smaller than most cross-sectional

studies (N<30) (P. van Eijndhoven et al. 2016; Gryglewski et al. 2018; Koenig et al. 2018; J. L. Phillips et al. 2015; Pirnia et al. 2016; Sartorius et al. 2016; Suh et al. 2019a). A recent consortium study has addressed several of these gaps in the literature, including a large MDD sample for sufficient power and a longitudinal design within a treatment paradigm that includes a placebo arm (Bartlett et al. 2018). Examining average values from 5 a priori selected regions (rostral and caudal anterior cingulate cortex (ACC), lateral orbitofrontal cortex, rostral middle frontal cortex and hippocampus), they found that only early thickening in the rostral ACC during the first week of treatment was associated with SSRI response at week 8 (Bartlett et al. 2018). No significant associations between pre-treatment cortical thickness and week-8 response were observed. This study, however, did not employ a whole-brain approach to capture information on cortical thickness in regions outside the pre-selected regions of interest (ROI).

We used a vertex/surface-based method to calculate cortical thickness with FreeSurfer, which utilizes a triangulated mesh to model the two surfaces that delineate the cerebral cortex: the pial boundary separating GM and cerebrospinal fluid (CSF) and the WM boundary that lies below cortical GM. There are several advantages to vertex-based methods when compared to conventional voxel-based morphometry, including sub-voxel accuracy, topological continuity, robustness to varying acquisition and scanner parameters and decreased susceptibility to partial volume effects (Clarkson et al. 2011; Fischl 2012).

Multi-site, multi-scanner effects are known to be a complex issue among the increasing number of large multi-site neuroimaging studies, which are necessary for

increasing sample sizes conducive to reliably detecting effects (Jahanshad et al. 2019; Hawco et al. 2018; Tozzi et al. undefined/ed). In recent years, the ComBat algorithm (Fortin et al. 2017; Johnson, Li, and Rabinovic 2007), originally developed for correcting batch effects in genomics, has been applied to correcting site- and scanner-associated variation (Bartlett et al. 2018; Yu et al. 2018). Here we have taken a similar approach, applying this algorithm to vertex-wise datapoints in our sample to accommodate our whole-brain analyses.

We present a study novel in its simultaneous whole-brain approach, large sample of MDD and HC participants and 16-week sequential treatment designed to investigate associations with antidepressant response. Our primary objectives were to investigate vertex-wise pre-treatment features of cortical thickness associated with antidepressant response at 8 weeks and 16 weeks of treatment. Specifically, we aimed to determine whether MDD participants who achieved clinical response at 8 and/or 16 weeks had differences in cortical thickness at baseline compared to those who did not achieve clinical response. We also aimed to identify differences between MDD and HC participants at baseline. We hypothesized that thinner cortex at baseline, which has been associated with increased vulnerability to MDD (Hao et al. 2017; Papmeyer et al. 2015), would be associated with worse response to treatment. A final aim of the study was to determine whether there were measurable changes in cortical thickness over the 8-week course of treatment with escitalopram. It has been suggested that antidepressant treatment may cause thickening of the cortex (Koenig et al. 2018; J. L. Phillips et al. 2015), but this hypothesis has not been confirmed in larger trials or in whole-brain analyses and is

contradicted by preclinical findings, in which stress-related decreases in cortical thickness are not normalized following SSRI administration (Lyttle et al. 2015). To our knowledge, this is the first treatment study to examine whole-brain group differences and longitudinal changes in cortical thickness with a sample size that can support this statistical approach (Pardoe, Abbott, and Jackson 2013). This last point in particular, combined with a reliable site correction method, is key to interrogating the often contradictory and inconclusive cortical thickness findings in recent years (Suh et al. 2019b).

4.2 Materials and Methods

Full details of the clinical, neuroimaging and biomarker protocols and patient outcomes are available elsewhere (Kennedy et al. 2019; Lam et al. 2016; MacQueen et al. 2019). The protocol was approved by the Research Ethics Boards at each institution. Information pertaining to inclusion/exclusion criteria and MRI acquisition parameters for this cohort can be found in Supplemental Information. The CONSORT diagram outlining the flow of participants throughout the 16-week clinical trial can be found in (Kennedy et al. 2019). MDD participants were aged 18-60 meeting DSM-IV-TR criteria for a major depressive episode (duration > 3 months) and HC subjects were matched for age and sex.

4.2.1 Treatment protocol

At the baseline visit (week 0), all MDD participants started treatment with escitalopram 10mg daily, flexible-dosage, with a maximum dose of 20mg per day. At week 8, responders (defined as a greater than 50% reduction in MADRS score (Lam et al.

2016)) continued to receive escitalopram for a further eight weeks. Participants who did

not respond received aripiprazole at 2-10mg per day (flexible-dosage), a first-line

adjunctive agent chosen based on clinical guidelines set out by the Canadian Network for

Mood and Anxiety Treatments (CANMAT) (Kennedy et al. 2016).

4.2.2 Imaging processing for cortical thickness analysis

We obtained T1-weighted images at weeks 0 (baseline), 2 and 8. Raw images

were pre-processed using the fully automated pipeline from FreeSurfer (version 6.0)

(https://surfer.nmr.mgh.harvard.edu/) (Fischl 2012). After motion correction and

averaging (Reuter, Rosas, and Fischl 2010), surrounding non-brain tissue is removed

(Ségonne et al. 2004) and the images then undergo a transformation to standard Talairach

space and intensity normalization (Sled, Zijdenbos, and Evans 1998). The boundary

between GM and WM undergoes tessellation to a triangular mesh and topological

corrections are made. The GM/WM and pial surfaces are then deformed to certain

locations where the greatest shifts in intensity occur and which indicate boundaries

between different tissue compartments (WM/GM/cerebrospinal fluid) (Dale, Fischl, and

Sereno 1999). These surfaces are then inflated to a spherical model and registered to the

MNI atlas. For each participant, cortical thickness values are measured for each vertex on

the cortical surface mesh as the shortest distance between the reconstructed pial and

GM/WM surfaces.

For longitudinal analyses, all images completed the FreeSurfer two-stage

longitudinal processing stream (Reuter et al. 2012). The first step is the creation of a

"base" unbiased template for each subject based on images from all timepoints (Reuter, Rosas, and Fischl 2010). Cross-sectional images are then processed for longitudinal analysis using the base template for initialization for skull stripping, transformation and registration, in order to minimize random error and preserve stable within-subject features across all timepoints (Reuter et al. 2012). First-pass quality assurance following all FreeSurfer pre-processing involved checking for correct skull strips and registration to Talairach space.

Quality assurance protocols from the Enhancing Neuroimaging Genetics through Meta-analysis (ENIGMA) consortium (http://enigma.ini.usc.edu/; April 2017) were used to assess results from the automatic cortical segmentation procedure outlined above. These protocols comprise an outlier detection analysis for parcellated ROIs as per the Desikan-Killiany atlas, visual inspection of internal segmentation (using sampled coronal and axial slices) and external surface reconstruction for each subject. Participants whose images failed at least one of these three steps were flagged for further inspection and manual edits of the main structural volume and white matter volume were made as necessary (254/795 total images over 3 timepoints). Edits were confined to cleaning up the pial boundary and on the white matter mask in the temporal lobes to improve segmentation of the gray matter from CSF on the superior aspect and gray matter from white matter in the temporal regions of the cortex, respectively. The resulting cortical maps were smoothed with a Gaussian kernel of 15mm full width at half-maximum in preparation for statistical group analyses.

4.2.3 Site effect corrections

Following FreeSurfer pre-processing and the completion of final quality checks, the Python implementation of the ComBat algorithm was used to correct for site effects on a vertex-wise basis (GitHub repository: https://github.com/ncullen93/neuroCombat). Full details can be found in Supplementary Materials.

4.2.4 Statistical analyses

All statistical analyses for demographics were performed in the open-source software Python (version 3.6) using the *scipy* library. MDD and HC groups were compared in terms of age, years of education and baseline MADRS using the Student's t-test. Age, years of education, age of illness onset, MADRS scores at the three timepoints, number of previous major depressive episodes and duration of illness were compared between the response groups using analysis of variance. Proportions of females/males were compared using a chi-square test (see Table 1).

At week 8, MDD participants were either escitalopram responders ('ESC-8') or non-responders ('NR-8'). At the week 16 endpoint, there were three groups based on response: those who continued to respond to escitalopram throughout the 16 weeks ('ESC-16'), participants who responded to adjunctive aripiprazole ('ARI-16') and patients who continued to be non-responders by week 16 despite the addition of aripiprazole ('NR-16'). In a secondary analysis, ESC-16 and ARI-16 groups were combined to define overall response at week 16. A small sample of five MDD

participants who were in the ESC-8 group but were no longer responding by week 16 were excluded only in subsequent analyses (Figure 1).

To test for changes in vertex-wise cortical thickness as a result of antidepressant treatment over 8 weeks, longitudinal analyses were performed using repeated-measures ANCOVA for the following groups to test for any changes, controlled for sex, age and age^2: the HC group, the MDD group and the three week-16 response subgroups. We also tested for differences in longitudinal changes from baseline to week 8 between ESC-8 and NR-8 groups to assess any between-group effects of escitalopram treatment. We measured longitudinal change using symmetrized percent change (SPC), the rate of change with respect to average thickness over the two timepoints. Using the average thickness renders SPC particularly robust to noise effects (Reuter et al. 2012; Tamnes et al. 2017).

Cross-sectional analyses consisted of whole-brain, vertex-wise comparisons between the aforementioned groups at baseline, and pairs of timepoints were used for longitudinal analyses (baseline to week 2, week 2 to week 8, baseline to week 8). General linear modelling (ANCOVA) was used for both between- and within-group analyses using FreeSurfer statistical tools, incorporating age, age^2 and sex as covariates (as per Perlman et al. 2017; Bartlett et al. 2018).

All vertex-wise results were corrected for multiple comparisons and separate hemisphere testing using Monte Carlo simulation (10,000 iterations) with a vertex-wise threshold of p=0.01 and cluster thresholding at p=0.05 (Hagler, Saygin, and Sereno 2006). Briefly, this technique involves indexing based on a lookup table provided by

FreeSurfer that has tabulated p-values corresponding to various cluster sizes at different smoothing levels. These values are derived from Gaussian Monte Carlo simulations of a z-field synthesized onto the cortical surface that is then thresholded to extract the largest cluster at a range of p-value thresholds, repeated 10,000 times. An additional Bonferroni correction was applied to take into account multiple exploratory cross-sectional contrasts (9 in total) for a post-correction cluster p-value threshold of p=0.0056 (family-wise threshold of 0.05 divided by the number of tests), set as the display threshold for all figures.

To examine in more detail how cortical thickness is related to the extent of improvement following 8 and 16 weeks of pharmacotherapy, we tested the relationship between percent change in MADRS score at 8 and 16 weeks and baseline cortical thickness within the MDD group using multiple linear regression (controlling for sex, age and age^2). Significant clusters from the whole-brain analysis were chosen as regions of interest, and a mask was created to extract average thickness values for each participant. Scatterplots were constructed to display the distributions and curves of best fit for each group in R version 3.4.1 (https://www.r-project.org/), using the package *ggplot2*.

4.3 Results

Baseline neuroimaging data were available for a total of 308 participants. Following the systematic quality screen as described above, we obtained FreeSurfer outputs for baseline images of usable quality from 181 MDD and 95 HC participants. 32 images were excluded on the basis of poor overall segmentation that could not be

manually corrected. There were no significant differences between MDD and HC on mean age or female:male ratio, although there was a significant difference in years of education (t=5.67; *p*=0.000) and baseline MADRS score (t=-63.88; p=0.000) (Table 1). By week 8, 22 participants had dropped out to give a sample size of 159 MDD participants, with a further reduction to 141 participants at week 16 (see Figure 1 for final MDD subgroup sample sizes). As indicated by omnibus tests, differences in age, sex, years of education, age of onset, illness duration, baseline severity and number of previous episodes were not significant between week-16 subgroups (Supplementary Table 1).

4.3.1 Baseline cross-sectional analyses between groups

Whole-brain analyses revealed that the MDD group exhibited thinner cortex at baseline in the left rostral middle frontal cortex (Figure 2; Table 2). There were no statistically significant differences at baseline in cortical thickness across the whole brain between the ESC-8 and NR-8 groups, between the ESC-16, ARI-16 and NR-16 groups, nor between the combined week-16 response group (ESC-16+ARI-16) and NR-16.

4.3.2 8-week longitudinal changes in cortical thickness

There were no significant longitudinal changes in cortical thickness within the HC group, the pooled MDD group or week-16 subgroups from baseline to week 8. Trajectories of cortical thickness change over the course of escitalopram treatment between ESC-8 and NR-8 groups were not found to be different.

4.3.3 Relationship between baseline cortical thickness and improvement in symptom severity

There were no significant relationships between baseline cortical thickness and extent of treatment response (percent improvement in MADRS scores) at either 8 or 16 weeks within the MDD group.

4.4 Discussion

In a large sample, with the novel combination of robust site effect correction and a vertex-based whole-brain approach, we found no significant differences at baseline between week-8 or week-16 responders and non-responders. Therefore, we were unable to confirm the hypothesis that non-responders to pharmacotherapy would exhibit thinner cortex at baseline and subsequently, the notion that cortical thickness might be a useful biomarker for treatment response. We found that the MDD group exhibited thinner cortex in the left rostral middle frontal (RMF) cortex as compared to HC. This result is a replication of findings from previous studies. Two studies found thinner cortex in the RMF region in MDD compared to HC, although one found a bilateral effect (Zhao et al. 2017) whereas the other also observed thinning in only the left region (Peng et al. 2015). Abnormalities in structural asymmetry in the RMF cortex has been reported in treatment-naïve MDD (Zuo et al. 2019). We note that these three studies also found additional effects of thinner cortex in other regions in MDD that we did not replicate. Additionally, increased thickness in the RMF region over the course of treatment has been found to be indicative of remission in two separate studies (J. L. Phillips et al. 2015; Saricicek

Aydogan et al. 2019)-- however, in the current study we did not find significant changes

in thickness within any response group on a whole-brain basis. Interestingly, in one study

thinner bilateral RMF at baseline was correlated to better response in the placebo group

(Bartlett et al. 2018) and *increased* thickness in this region was negatively correlated to

symptom severity in MDD (Qiu et al. 2014). It appears that even for one region that is

implicated relatively often in the literature, previous reports are mixed and are often

accompanied by other findings that have not been replicated.

Although our primary analyses yielded largely null results as far as the association

between cortical thickness and antidepressant response, they are fairly consistent with

recent studies that have been published on the topic of cortical thickness as a

neuroimaging biomarker in MDD. Despite some positive preliminary results, pre-

treatment cortical thickness predictors of treatment response that are robust and replicable

have yet to be discovered (J. L. Phillips et al. 2015; Suh et al. 2019a; Bartlett et al. 2018).

Similarly, studies on diagnostic group differences in cortical thickness, although more

numerous, have yet to converge on a set of replicable differences (Suh et al. 2019b),

finding both thinner and thicker regions among disparate regions in MDD. The largest

cortical thickness analysis to date from the ENIGMA consortium, with 1902 MDD

patients and 7658 HC participants drawn from 20 cohorts, found small absolute effect

sizes of thinning in MDD (Cohen's *d* no larger than 0.14) despite its robust statistical

power. Studies with sample sizes ranging from 50 to >100 per group tend towards null

results for statistical testing of the MDD vs HC comparison (Perlman et al. 2017;

Saricicek Aydogan et al. 2019). This indicates that smaller samples, particularly in the

context of cortical thickness analysis (Button et al. 2013), may run an increased risk of false positive results. Inconsistent methods for multiple comparison corrections and variable utilization of region-based vs. vertex-wise approaches have also contributed to the considerable heterogeneity among studies. Other clinical characteristics of the MDD sample may also influence the final results, such as varying definitions of response/remission, medication status and disease chronicity.

When the extant literature and the current results are taken together, it seems unlikely that cortical thickness alone could be used as a reliable predictor of short-term treatment response to pharmacotherapy. However, it has been shown to be useful for mapping and predicting clinical response to electroconvulsive therapy in a regionally consistent manner in several studies (P. van Eijndhoven et al. 2016; Pirnia et al. 2016; Sartorius et al. 2016; Schmitgen et al. 2019; Wade et al. 2017). Effectively, cortical thickness represents a totality of numerous microscopic properties; shown to be a relatively stable measure over the lifespan (Storsve et al. 2014; Hogstrom et al. 2013), it appears likely that any subtle structural characteristics predictive of the extent of short-term response to pharmacotherapy, if they exist, are not well-reflected in this measure.

We found no significant longitudinal changes in cortical thickness following 8 weeks of escitalopram treatment, nor any differences between responders and non-responders. Most studies reporting longitudinal changes in cortical thickness (Bartlett et al. 2018; Koenig et al. 2018; J. L. Phillips et al. 2015) have focused on pre-determined regions of interest for analysis, lending them the advantage of increased statistical power. Other studies have found no change in either cortical volume nor thickness within several

weeks of SSRI or SNRI treatment (Fu et al. 2015; Lyttle et al. 2015). It is also conceivable that 8 weeks might be too short a period of time to detect cortical thickness changes via MRI.

This study addressed a gap in the literature regarding identification of baseline features associated with differential treatment response (M. L. Phillips et al. 2015), particularly being the first neuroimaging study to incorporate sequential adjunctive treatment following lack of response to a first-line antidepressant medication (Kennedy et al. 2016). Given the variety of cortical regions that have been identified as being altered or abnormal in MDD, we opted for a whole-brain, exploratory approach, taking advantage of the relatively large sample size afforded by the multi-site effort of the CAN-BIND trial. Other advantages of the current study include controlling for heterogeneity introduced by different MRI scanners and varying acquisition parameters by correcting for these effects on a vertex-wise basis using the ComBat algorithm. We also carried out blinded systematic manual correction of FreeSurfer outputs, which has been shown to provide more accurate segmentations when compared to uncorrected outputs (Popescu et al. 2016).

There are also several limitations associated with the study. First, although we have a relatively large sample size overall for both the MDD and HC groups, this advantage is reduced once we subdivide the MDD group based on week-16 response. Second, our age range was relatively large, possibly obscuring potential age-related trajectories and mechanisms of disease progression. Although we attempted to mitigate this limitation by controlling for both age and age^2 in all neuroimaging analyses, it is

possible that age-specific relationships may emerge in studies with narrower age ranges or greater sample sizes. Third, we excluded subcortical and cerebellar regions in the analysis, as we focused only on cerebral cortical thickness, and therefore could not ascertain any longitudinal changes or baseline associations with treatment response in these areas. Moreover, we did not control for the potential confounding variable of cigarette smoking status (Zorlu et al. 2017). Finally, mass univariate analyses are not sufficient to discriminate structural features that could aid in informing individualized treatment.

In conclusion, we show that cortical thickness in MDD at baseline was not associated to antidepressant response at 8 or 16 weeks, nor was it shown to change over an 8-week period of escitalopram treatment. We did replicate previous findings of cortical thinning in the left RMF cortex in MDD. Future studies might investigate not only univariate approaches to isolating potential biomarkers, but also multivariate methods incorporating multiple measures, an approach requiring a sufficient number of participants and clearly defined patient samples (Kim and Na 2018; Raamana et al. 2014). Another promising approach is the extraction of advanced multi-variate network-level features from whole-brain thickness features to assess their utility in predicting response to treatment (Raamana and Strother 2018). The emergence of larger consortia with sufficient power to identify subgroups based on biomarkers is a promising sign for the field (Brunoni et al. 2015; Lam et al. 2016; Schmaal et al. 2017; Trivedi et al. 2016). Methods range from retrospective grouping of subjects based on some outcome parameter

(as shown here with 'response') to using unsupervised, data-driven machine learning algorithms to model underlying patterns of variability (Drysdale et al. 2017).

4.5 Funding & Disclosure

4.6 References

Bartlett, Elizabeth A., Christine DeLorenzo, Priya Sharma, Jie Yang, Mengru Zhang, Eva

    Petkova, Myrna Weissman, et al. 2018. "Pretreatment and Early-Treatment

    Cortical Thickness Is Associated with SSRI Treatment Response in Major

    Depressive Disorder." Neuropsychopharmacology: Official Publication of the

    American College of Neuropsychopharmacology 43 (11): 2221–30.

    https://doi.org/10.1038/s41386-018-0122-9.

Brunoni, Andre Russowsky, Bernardo Sampaio-Junior, Adriano Henrique Moffa, Lucas

    Borrione, Barbara Schwair Nogueira, Luana Vanessa Marotti Aparicio, Beatriz

    Veronezi, et al. 2015. "The Escitalopram versus Electric Current Therapy for

    Treating Depression Clinical Study (ELECT-TDCS): Rationale and Study

    Design of a Non-Inferiority, Triple-Arm, Placebo-Controlled Clinical Trial." Sao

    Paulo Medical Journal 133 (3): 252–63. https://doi.org/10.1590/1516-

    3180.2014.00351712.

Button, Katherine S., John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan

    Flint, Emma S. J. Robinson, and Marcus R. Munafò. 2013. "Power Failure: Why

    Small Sample Size Undermines the Reliability of Neuroscience." Nature

    Reviews Neuroscience 14 (5): 365–76. https://doi.org/10.1038/nrn3475.

Cipriani, Andrea, Toshi A. Furukawa, Georgia Salanti, Anna Chaimani, Lauren Z.

    Atkinson, Yusuke Ogawa, Stefan Leucht, et al. 2018. "Comparative Efficacy and

    Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults

    with Major Depressive Disorder: A Systematic Review and Network Meta-

Analysis." The Lancet 391 (10128): 1357–66. https://doi.org/10.1016/S0140-6736(17)32802-7.

Clarkson, Matthew J., M. Jorge Cardoso, Gerard R. Ridgway, Marc Modat, Kelvin K. Leung, Jonathan D. Rohrer, Nick C. Fox, and Sébastien Ourselin. 2011. "A Comparison of Voxel and Surface Based Cortical Thickness Estimation Methods." Neuroimage 57 (3): 856–65. https://doi.org/10.1016/j.neuroimage.2011.05.053.

Cremers, Henk R., Tor D. Wager, and Tal Yarkoni. 2017. "The Relation between Statistical Power and Inference in FMRI." PloS One 12 (11): e0184923. https://doi.org/10.1371/journal.pone.0184923.

Dale, A. M., B. Fischl, and M. I. Sereno. 1999. "Cortical Surface-Based Analysis - I. Segmentation and Surface Reconstruction." NeuroImage 9 (2): 179–94. https://doi.org/10.1006/nimg.1998.0395.

Dowlati, Yekta, Nathan Herrmann, Walter Swardfager, Helena Liu, Lauren Sham, Elyse K. Reim, and Krista L. Lanctot. 2010. "A Meta-Analysis of Cytokines in Major Depression." Biological Psychiatry 67 (5): 446–57. https://doi.org/10.1016/j.biopsych.2009.09.033.

Drysdale, Andrew T., Logan Grosenick, Jonathan Downar, Katharine Dunlop, Farrokh Mansouri, Yue Meng, Robert N. Fetcho, et al. 2017. "Resting-State Connectivity Biomarkers Define Neurophysiological Subtypes of Depression." Nature Medicine 23 (1): 28–38. https://doi.org/10.1038/nm.4246.

Eijndhoven, P. van, P. Mulders, L. Kwekkeboom, I. van Oostrom, M. van Beek, J. Janzing, A. Schene, and I. Tendolkar. 2016. "Bilateral ECT Induces Bilateral Increases in Regional Cortical Thickness." Translational Psychiatry 6 (August): e874. https://doi.org/10.1038/tp.2016.139.

Eijndhoven, Philip van, Guido van Wingen, Maartje Katzenbauer, Wouter Groen, Ralf Tepest, Guillen Fernandez, Jan Buitelaar, and Indira Tendolkar. 2013. "Paralimbic Cortical Thickness in First-Episode Depression: Evidence for Trait-Related Differences in Mood Regulation." American Journal of Psychiatry 170 (12): 1477–86. https://doi.org/10.1176/appi.ajp.2013.12121504.

Fischl, Bruce. 2012. "FreeSurfer." NeuroImage 62 (2): 774–81. https://doi.org/10.1016/j.neuroimage.2012.01.021.

Fonseka, Bernice A., Natalia Jaworska, Allegra Courtright, Frank P. MacMaster, and Glenda M. MacQueen. 2016. "Cortical Thickness and Emotion Processing in Young Adults with Mild to Moderate Depression: A Preliminary Study." Bmc Psychiatry 16 (February): 38. https://doi.org/10.1186/s12888-016-0750-8.

Fortin, Jean-Philippe, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem Aselcioglu, Phil Adams, Crystal Cooper, et al. 2017. "Harmonization of Cortical Thickness Measurements across Scanners and Sites." BioRxiv, June. https://doi.org/10.1101/148502.

Fu, Cynthia HY, Sergi G. Costafreda, Anjali Sankar, Tracey M. Adams, Mark M. Rasenick, Peng Liu, Robert Donati, Luigi A. Maglanoc, Paul Horton, and Lauren B. Marangell. 2015. "Multimodal Functional and Structural Neuroimaging

Investigation of Major Depressive Disorder Following Treatment with Duloxetine." BMC Psychiatry 15 (1): 82. https://doi.org/10.1186/s12888-015-0457-2.

Gryglewski, Gregor, Pia Baldinger-Melich, René Seiger, Godber Mathis Godbersen, Paul Michenthaler, Manfred Klöbl, Benjamin Spurny, et al. 2018. "Structural Changes in Amygdala Nuclei, Hippocampal Subfields and Cortical Thickness Following Electroconvulsive Therapy in Treatment-Resistant Depression: Longitudinal Analysis." The British Journal of Psychiatry, 1–9. https://doi.org/10.1192/bjp.2018.224.

Hagler, Donald J., Ayse Pinar Saygin, and Martin I. Sereno. 2006. "Smoothing and Cluster Thresholding for Cortical Surface-Based Group Analysis of FMRI Data." NeuroImage 33 (4): 1093–1103. https://doi.org/10.1016/j.neuroimage.2006.07.036.

Hao, Xuejun, Ardesheer Talati, Stewart A. Shankman, Jun Liu, Jurgen Kaiser, Craig E. Tenke, Virginia Warner, et al. 2017. "Stability of Cortical Thinning in Persons at Increased Familial Risk for Major Depressive Disorder Across 8 Years." Biological Psychiatry. Cognitive Neuroscience and Neuroimaging 2 (7): 619–25. https://doi.org/10.1016/j.bpsc.2017.04.009.

Hawco, Colin, Joseph D. Viviano, Sofia Chavez, Erin W. Dickie, Navona Calarco, Peter Kochunov, Miklos Argyelan, et al. 2018. "A Longitudinal Human Phantom Reliability Study of Multi-Center T1-Weighted, DTI, and Resting State FMRI

Data." Psychiatry Research. Neuroimaging 282 (December): 134–42.

https://doi.org/10.1016/j.pscychresns.2018.06.004.

Hogstrom, Larson J., Lars T. Westlye, Kristine B. Walhovd, and Anders M. Fjell. 2013.

"The Structure of the Cerebral Cortex Across Adult Life: Age-Related Patterns

of Surface Area, Thickness, and Gyrification." Cerebral Cortex 23 (11): 2521–

30. https://doi.org/10.1093/cercor/bhs231.

Jahanshad, Neda, Joshua I. Faskowitz, Gennady Roshchupkin, Derrek Hibar, Boris A.

Gutman, Nicholas J. Tustison, Hieab H. H. Adams, et al. 2019. "MULTI-SITE

META-ANALYSIS OF MORPHOMETRY." IEEE/ACM Transactions on

Computational Biology and Bioinformatics, May.

https://doi.org/10.1109/TCBB.2019.2914905.

Johnson, W. Evan, Cheng Li, and Ariel Rabinovic. 2007. "Adjusting Batch Effects in

Microarray Expression Data Using Empirical Bayes Methods." Biostatistics

(Oxford, England) 8 (1): 118–27. https://doi.org/10.1093/biostatistics/kxj037.

Kennedy, Sidney H., and Amanda K. Ceniti. 2018. "Unpacking Major Depressive

Disorder: From Classification to Treatment Selection." Canadian Journal of

Psychiatry. Revue Canadienne de Psychiatrie 63 (5): 308–13.

https://doi.org/10.1177/0706743717748883.

Kennedy, Sidney H., Raymond W. Lam, Roger S. McIntyre, S. Valérie Tourjman, Venkat

Bhat, Pierre Blier, Mehrul Hasnain, et al. 2016. "Canadian Network for Mood

and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the

Management of Adults with Major Depressive Disorder: Section 3.

Pharmacological Treatments." The Canadian Journal of Psychiatry 61 (9): 540–60. https://doi.org/10.1177/0706743716659417.

Kennedy, Sidney H., Raymond W. Lam, Susan Rotzinger, Roumen V. Milev, Pierre Blier, Jonathan Downar, Kenneth R. Evans, Faranak Farzan, Jane A. Foster, and Benicio N. Frey. 2019. "Symptomatic and Functional Outcomes and Early Prediction of Response to Escitalopram Monotherapy and Sequential Adjunctive Aripiprazole Therapy in Patients With Major Depressive Disorder: A CAN-BIND-1 Report." The Journal of Clinical Psychiatry 80 (2): 0–0. https://doi.org/10.4088/JCP.18m12202.

Kim, Yong-Ku, and Kyoung-Sae Na. 2018. "Application of Machine Learning Classification for Structural Brain MRI in Mood Disorders: Critical Review from a Clinical Perspective." Progress in Neuro-Psychopharmacology & Biological Psychiatry 80 (January): 71–80. https://doi.org/10.1016/j.pnpbp.2017.06.024.

Koenig, Julian, Melinda Westlund Schreiner, Bonnie Klimes-Dougan, Benjamin Ubani, Bryon A. Mueller, Kelvin O. Lim, Michael Kaess, and Kathryn R. Cullen. 2018. "Increases in Orbitofrontal Cortex Thickness Following Antidepressant Treatment Are Associated with Changes in Resting State Autonomic Function in Adolescents with Major Depression – Preliminary Findings from a Pilot Study." Psychiatry Research: Neuroimaging 281 (November): 35–42. https://doi.org/10.1016/j.pscychresns.2018.08.013.

Lam, Raymond W., Roumen Milev, Susan Rotzinger, Ana C. Andreazza, Pierre Blier, Colleen Brenner, Zafiris J. Daskalakis, et al. 2016. "Discovering Biomarkers for

Antidepressant Response: Protocol from the Canadian Biomarker Integration

Network in Depression (CAN-BIND) and Clinical Characteristics of the First

Patient Cohort." Bmc Psychiatry 16 (April): 105.

https://doi.org/10.1186/s12888-016-0785-x.

Lyttle, Kerise, Yu Ohmura, Kohtarou Konno, Takayuki Yoshida, Takeshi Izumi,

Masahiko Watanabe, and Mitsuhiro Yoshioka. 2015. "Repeated Fluvoxamine

Treatment Recovers Juvenile Stress-Induced Morphological Changes and

Depressive-like Behavior in Rats." Brain Research 1616 (August): 88–100.

https://doi.org/10.1016/j.brainres.2015.04.058.

MacQueen, Glenda M., Stefanie Hassel, Stephen R. Arnott, Addington Jean, Christopher

R. Bowie, Signe L. Bray, Andrew D. Davis, et al. 2019. "The Canadian

Biomarker Integration Network in Depression (CAN-BIND): Magnetic

Resonance Imaging Protocols." Journal of Psychiatry & Neuroscience: JPN 44

(3): 1–14. https://doi.org/10.1503/jpn.180036.

Narr, Katherine L., Roger P. Woods, Paul M. Thompson, Philip Szeszko, Delbert

Robinson, Teodora Dimtcheva, Mala Gurbani, Arthur W. Toga, and Robert M.

Bilder. 2007. "Relationships between IQ and Regional Cortical Gray Matter

Thickness in Healthy Adults." Cerebral Cortex 17 (9): 2163–71.

https://doi.org/10.1093/cercor/bhl125.

Papmeyer, Martina, Stephen Giles, Jessica E. Sussmann, Shauna Kielty, Tiffany Stewart,

Stephen M. Lawrie, Heather C. Whalley, and Andrew M. McIntosh. 2015.

"Cortical Thickness in Individuals at High Familial Risk of Mood Disorders as

They Develop Major Depressive Disorder." Biological Psychiatry 78 (1): 58–66. https://doi.org/10.1016/j.biopsych.2014.10.018.

Pardoe, Heath R., David F. Abbott, and Graeme D. Jackson. 2013. "Sample Size Estimates for Well-Powered Cross-Sectional Cortical Thickness Studies." Human Brain Mapping 34 (11). https://doi.org/10.1002/hbm.22120.

Paris, Joel. 2014. "The Mistreatment of Major Depressive Disorder." Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie 59 (3): 148–51. https://doi.org/10.1177/070674371405900306.

Peng, Daihui, Feng Shi, Gang Li, Drew Fralick, Ting Shen, Meihui Qiu, Jun Liu, Kaida Jiang, Dinggang Shen, and Yiru Fang. 2015. "Surface Vulnerability of Cerebral Cortex to Major Depressive Disorder." Plos One 10 (3): e0120704. https://doi.org/10.1371/journal.pone.0120704.

Perlman, Greg, Elizabeth Bartlett, Christine DeLorenzo, Myrna Weissman, Patrick McGrath, Todd Ogden, Tony Jin, et al. 2017. "Cortical Thickness Is Not Associated with Current Depression in a Clinical Treatment Study." Human Brain Mapping 38 (9): 4370–85. https://doi.org/10.1002/hbm.23664.

Phillips, Jennifer L., Lisa A. Batten, Philippe Tremblay, Fahad Aldosary, and Pierre Blier. 2015. "A Prospective, Longitudinal Study of the Effect of Remission on Cortical Thickness and Hippocampal Volume in Patients with Treatment-Resistant Depression." International Journal of Neuropsychopharmacology 18 (8). https://doi.org/10.1093/ijnp/pyv037.

Phillips, Mary L., Henry W. Chase, Yvette I. Sheline, Amit Etkin, Jorge R.C. Almeida, Thilo Deckersbach, and Madhukar H. Trivedi. 2015. "Identifying Predictors, Moderators, and Mediators of Antidepressant Response in Major Depressive Disorder: Neuroimaging Approaches." The American Journal of Psychiatry 172 (2): 124–38. https://doi.org/10.1176/appi.ajp.2014.14010076.

Pink, Anna, Scott A. Przybelski, Janina Krell-Roesch, Gorazd B. Stokin, Rosebud O. Roberts, Michelle M. Mielke, David S. Knopman, Clifford R. Jack, Ronald C. Petersen, and Yonas E. Geda. 2017. "Cortical Thickness and Depressive Symptoms in Cognitively Normal Individuals: The Mayo Clinic Study of Aging." Journal of Alzheimer's Disease: JAD 58 (4): 1273–81. https://doi.org/10.3233/JAD-170041.

Pirnia, T., S. H. Joshi, A. M. Leaver, M. Vasavada, S. Njau, R. P. Woods, R. Espinoza, and K. L. Narr. 2016. "Electroconvulsive Therapy and Structural Neuroplasticity in Neocortical, Limbic and Paralimbic Cortex." Translational Psychiatry 6 (June): e832. https://doi.org/10.1038/tp.2016.102.

Popescu, Veronica, Roel Klaver, Adriaan Versteeg, Pieter Voorn, Jos W. R. Twisk, Frederik Barkhof, Jeroen J. G. Geurts, and Hugo Vrenken. 2016. "Postmortem Validation of MRI Cortical Volume Measurements in MS." Human Brain Mapping 37 (6): 2223–33. https://doi.org/10.1002/hbm.23168.

Qiu, L., S. Lui, W. Kuang, X. Huang, J. Li, J. Li, J. Zhang, H. Chen, J. A. Sweeney, and Q. Gong. 2014. "Regional Increases of Cortical Thickness in Untreated, First-

Episode Major Depressive Disorder." Translational Psychiatry 4 (April): e378. https://doi.org/10.1038/tp.2014.18.

Raamana, PR, and Stephen C. Strother. 2018. "Graynet: Single-Subject Morphometric Networks for Neuroscience Connectivity Applications." Journal of Open Source Software 3 (30): 924. https://doi.org/10.21105/joss.00924.

Raamana, Pradeep Reddy, Wei Wen, Nicole A. Kochan, Henry Brodaty, Perminder S. Sachdev, Lei Wang, and Mirza Faisal Beg. 2014. "The Sub-Classification of Amnestic Mild Cognitive Impairment Using MRI-Based Cortical Thickness Measures." Frontiers in Neurology 5: 76. https://doi.org/10.3389/fneur.2014.00076.

Reuter, Martin, H. Diana Rosas, and Bruce Fischl. 2010. "Highly Accurate Inverse Consistent Registration: A Robust Approach." NeuroImage 53 (4): 1181–96. https://doi.org/10.1016/j.neuroimage.2010.07.020.

Reuter, Martin, Nicholas J. Schmansky, H. Diana Rosas, and Bruce Fischl. 2012. "Within-Subject Template Estimation for Unbiased Longitudinal Image Analysis." NeuroImage 61 (4): 1402–18. https://doi.org/10.1016/j.neuroimage.2012.02.084.

Saricicek Aydogan, A., E. Oztekin, M. E. Esen, S. Dusmez, F. Gelal, L. Besiroğlu, and N. Zorlu. 2019. "Cortical Thickening in Remitters Compared to Non-remitters with Major Depressive Disorder Following 8-week Antidepressant Treatment." Acta Psychiatrica Scandinavica 140 (3): 217–26. https://doi.org/10.1111/acps.13065.

Sartorius, Alexander, Traute Demirakca, Andreas Böhringer, Christian Clemm von Hohenberg, Suna Su Aksay, Jan Malte Bumb, Laura Kranaster, and Gabriele Ende. 2016. "Electroconvulsive Therapy Increases Temporal Gray Matter Volume and Cortical Thickness." European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology 26 (3): 506–17. https://doi.org/10.1016/j.euroneuro.2015.12.036.

Schmaal, L., D. P. Hibar, P. G. Saemann, G. B. Hall, B. T. Baune, N. Jahanshad, J. W. Cheung, et al. 2017. "Cortical Abnormalities in Adults and Adolescents with Major Depression Based on Brain Scans from 20 Cohorts Worldwide in the ENIGMA Major Depressive Disorder Working Group." Molecular Psychiatry 22 (6): 900–909. https://doi.org/10.1038/mp.2016.60.

Schmitgen, Mike M., Katharina M. Kubera, Malte S. Depping, Henrike M. Nolte, Dusan Hirjak, Stefan Hofer, Julia H. Hasenkamp, et al. 2019. "Exploring Cortical Predictors of Clinical Response to Electroconvulsive Therapy in Major Depression." European Archives of Psychiatry and Clinical Neuroscience, July. https://doi.org/10.1007/s00406-019-01033-w.

Schüz, A., and G. Palm. 1989. "Density of Neurons and Synapses in the Cerebral Cortex of the Mouse." The Journal of Comparative Neurology 286 (4): 442–55. https://doi.org/10.1002/cne.902860404.

Ségonne, F., A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl. 2004. "A Hybrid Approach to the Skull Stripping Problem in MRI." NeuroImage 22 (3): 1060–75. https://doi.org/10.1016/j.neuroimage.2004.03.032.

Seldon, H. Lee. 2007. "Extended Neocortical Maturation Time Encompasses Speciation,

    Fatty Acid and Lateralization Theories of the Evolution of Schizophrenia and

    Creativity." Medical Hypotheses 69 (5): 1085–89.

    https://doi.org/10.1016/j.mehy.2007.03.001.

Shaw, Philip, Noor J. Kabani, Jason P. Lerch, Kristen Eckstrand, Rhoshel Lenroot, Nitin

    Gogtay, Deanna Greenstein, et al. 2008. "Neurodevelopmental Trajectories of

    the Human Cerebral Cortex." Journal of Neuroscience 28 (14): 3586–94.

    https://doi.org/10.1523/JNEUROSCI.5309-07.2008.

Sled, J. G., A. P. Zijdenbos, and A. C. Evans. 1998. "A Nonparametric Method for

    Automatic Correction of Intensity Nonuniformity in MRI Data." IEEE

    Transactions on Medical Imaging 17 (1): 87–97.

    https://doi.org/10.1109/42.668698.

Storsve, Andreas B., Anders M. Fjell, Christian K. Tamnes, Lars T. Westlye, Knut

    Overbye, Hilde W. Aasland, and Kristine B. Walhovd. 2014. "Differential

    Longitudinal Changes in Cortical Thickness, Surface Area and Volume across

    the Adult Life Span: Regions of Accelerating and Decelerating Change." The

    Journal of Neuroscience: The Official Journal of the Society for Neuroscience 34

    (25): 8488–98. https://doi.org/10.1523/JNEUROSCI.0391-14.2014.

Suh, Jee Su, Luciano Minuzzi, Lauren E. Cudney, William Maich, Maha Eltayebani,

    Claudio N. Soares, and Benicio N. Frey. 2019a. "Cerebral Cortical Thickness

    after Treatment with Desvenlafaxine Succinate in Major Depressive Disorder."

    NeuroReport 30 (5): 378. https://doi.org/10.1097/WNR.0000000000001211.

Suh, Jee Su, Maiko Abel Schneider, Luciano Minuzzi, Glenda M. MacQueen, Stephen C. Strother, Sidney H. Kennedy, and Benicio N. Frey. 2019b. "Cortical Thickness in Major Depressive Disorder: A Systematic Review and Meta-Analysis." Progress in Neuro-Psychopharmacology and Biological Psychiatry 88 (January): 287–302. https://doi.org/10.1016/j.pnpbp.2018.08.008.

Tamnes, Christian K., Megan M. Herting, Anne-Lise Goddings, Rosa Meuwese, Sarah-Jayne Blackmore, Ronald E. Dahl, Berna Güroğlu, Armin Raznahan, Elizabeth R. Sowell, Eveline A. Crone, and Kathryn L. Mills. 2017. "Development of the Cerebral Cortex across Adolescence: A Multisample Study of Inter-Related Longitudinal Changes in Cortical Volume, Surface Area, and Thickness." Journal of Neuroscience, 37 (12): 3402-3412.

Tozzi, Leonardo, Lisa Garczarek, Deborah Janowitz, Dan J. Stein, Katharina Wittfeld, Henrik Dobrowolny, Jim Lagopoulos, et al. undefined/ed. "Interactive Impact of Childhood Maltreatment, Depression, and Age on Cortical Brain Structure: Mega-Analytic Findings from a Large Multi-Site Cohort." Psychological Medicine, 1–12. https://doi.org/10.1017/S003329171900093X.

Trivedi, Madhukar H., Patrick J. McGrath, Maurizio Fava, Ramin V. Parsey, Benji T. Kurian, Mary L. Phillips, Maria A. Oquendo, et al. 2016. "Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC): Rationale and Design." Journal of Psychiatric Research 78 (July): 11–23. https://doi.org/10.1016/j.jpsychires.2016.03.001.

Wade, Benjamin S. C., Jing Sui, Gerhard Hellemann, Amber M. Leaver, Randall T.
Espinoza, Roger P. Woods, Christopher C. Abbott, Shantanu H. Joshi, and
Katherine L. Narr. 2017. "Inter and Intra-Hemispheric Structural Imaging
Markers Predict Depression Relapse after Electroconvulsive Therapy: A
Multisite Study." Translational Psychiatry 7 (12): 1270.
https://doi.org/10.1038/s41398-017-0020-7.

"WHO | Depression." n.d. WHO. Accessed December 12, 2018.
http://www.who.int/mental_health/management/depression/en/.

Yang, Xin-hua, Yi Wang, Jia Huang, Cui-ying Zhu, Xiao-qun Liu, Eric F. C. Cheung,
Guang-rong Xie, and Raymond C. K. Chan. 2015. "Increased Prefrontal and
Parietal Cortical Thickness Does Not Correlate with Anhedonia in Patients with
Untreated First-Episode Major Depressive Disorders." Psychiatry Research-
Neuroimaging 234 (1): 144–51.
https://doi.org/10.1016/j.pscychresns.2015.09.014.

Young, Juan Joseph, Tim Silber, Davide Bruno, Isaac Robert Galatzer-Levy, Nunzio
Pomara, and Charles Raymond Marmar. 2016. "Is There Progress? An Overview
of Selecting Biomarker Candidates for Major Depressive Disorder." Frontiers in
Psychiatry 7: 72. https://doi.org/10.3389/fpsyt.2016.00072.

Yu, Meichen, Kristin A. Linn, Philip A. Cook, Mary L. Phillips, Melvin McInnis,
Maurizio Fava, Madhukar H. Trivedi, Myrna M. Weissman, Russell T.
Shinohara, and Yvette I. Sheline. 2018. "Statistical Harmonization Corrects Site

Effects in Functional Connectivity Measurements from Multi-Site FMRI Data."

Human Brain Mapping 39 (11): 4213–27. https://doi.org/10.1002/hbm.24241.

Zhao, Youjin, Lizhou Chen, Wenjing Zhang, Yuan Xiao, Chandan Shah, Hongru Zhu, Minlan Yuan, et al. 2017. "Gray Matter Abnormalities in Non-Comorbid Medication-Naive Patients with Major Depressive Disorder or Social Anxiety Disorder." EBioMedicine 21 (July): 228. https://doi.org/10.1016/j.ebiom.2017.06.013.

Zorlu, Nabi, Vanessa Louise Cropley, Pelin Kurtgoz Zorlu, Dursun Hakan Delibas, Zehra Hilal Adibelli, Emel Pasa Baskin, Ozgur Sipahi Esen, Emre Bora, and Christos Pantelis. 2017. "Effects of Cigarette Smoking on Cortical Thickness in Major Depressive Disorder." Journal of Psychiatric Research 84 (January): 1–8. https://doi.org/10.1016/j.jpsychires.2016.09.009.

Zuo, Zhiwei, Shuhua Ran, Yao Wang, Chang Li, Qi Han, Qianying Tang, Wei Qu, and Haitao Li. 2019. "Asymmetry in Cortical Thickness and Subcortical Volume in Treatment-Naïve Major Depressive Disorder." NeuroImage : Clinical 21. https://doi.org/10.1016/j.nicl.2018.101614.

4.7 Figures and Tables
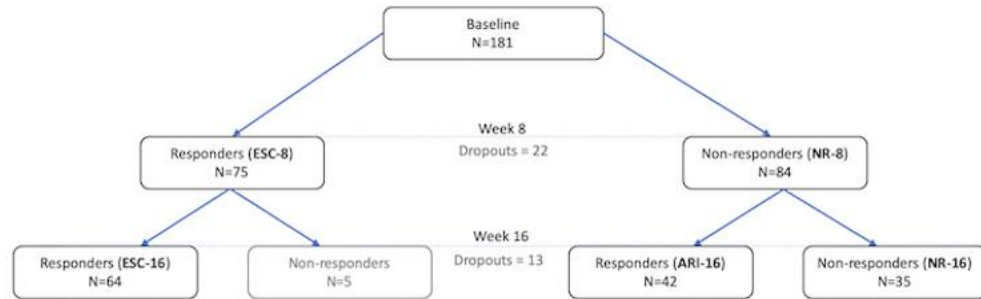
Figure 1. Number of participants in response subgroups at week 8 and week 16.
22 participants were lost to attrition at week 8 and an additional 13 had dropped out by week 16.
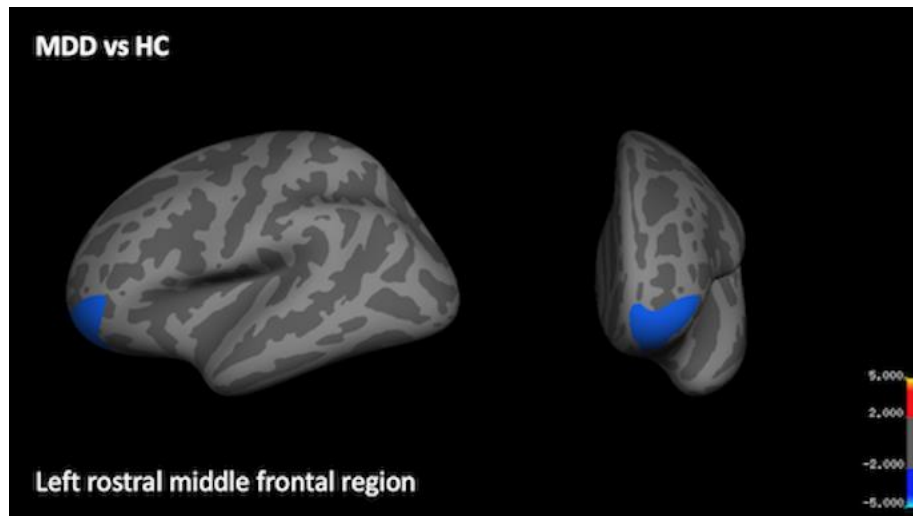
Figure 2. Region of thinner cortex at baseline in MDD as compared to HC.
A significant cluster indicating thinner cortex in the left rostral middle frontal region are displayed from the lateral and frontal views. Scale bar shows max -log(p) values, following corrections for multiple comparisons using Monte Carlo thresholding.
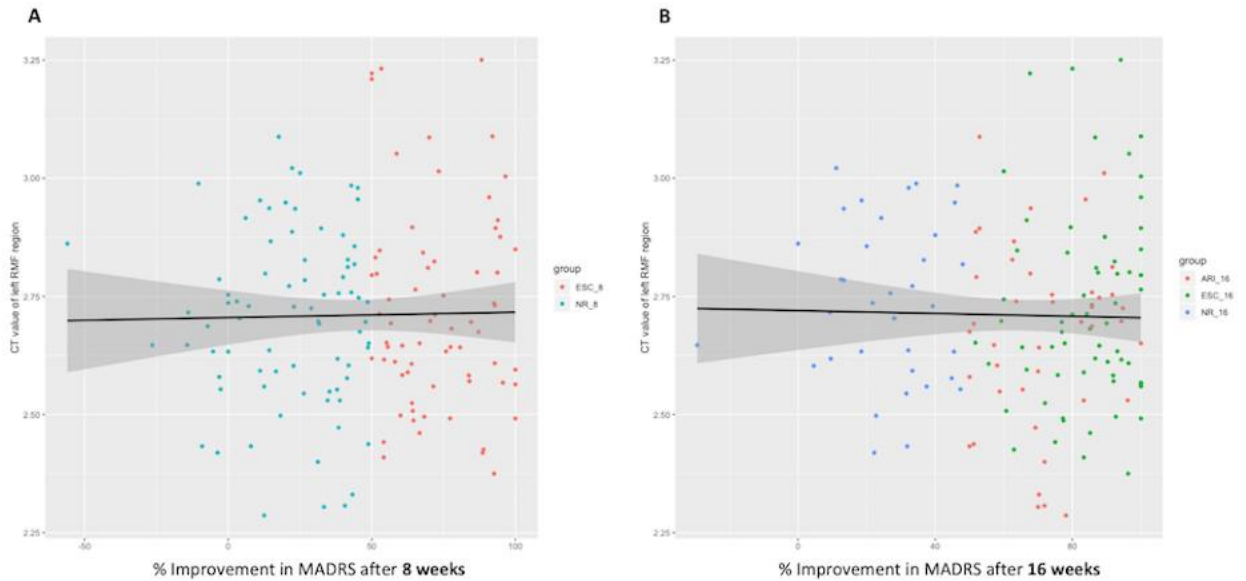
Figure 3. Scatterplots and regression lines depicting the relationship between RMF thickness and % improvement in MADRS scores, grouped by response.
Thickness values over the significant cluster in the RMF cortex at baseline were extracted for all MDD participants and plotted over % improvement in MADRS score at A) 8 weeks of escitalopram treatment and B) 16 weeks of escitalopram or adjunctive treatment with aripiprazole for each participant. Datapoints have been grouped by colour based on response group at the respective treatment timepoints.

Table 1. **Demographic and clinical information for the MDD and HC samples.**

	MDD Patients			Healthy Controls		
	N	Mean/ frequency	SD	*N*	Mean/ frequency	SD
Age [1]	181	35.03	12.56	95	32.99	10.87
Sex % F [1]	181	64.84	--	95	63.54	--
Years of Education [2]	181	16.87	2.08	95	18.40	2.23
Baseline MADRS [2]	181	29.84	5.65	95	0.82	1.73
Age of Illness Onset	175	20.38	10.26	--	--	--
Number of previous MDEs	132	4.02	2.64	--	--	--
% Change MADRS (8-week)	159	-45.79	32.43	--	--	--
% Change MADRS (16-week)	141	-65.71	27.70	--	--	--
% Responders @ 8 weeks	159	47.17	--	--	--	--
% Responders @ 16 weeks	141	75.20	--	--	--	--
% Family history of psychiatric illness	180	77.78	--	--	--	--

Superscripts indicate the significance of the test statistic comparing patient and healthy control samples. '1' – $p > 0.05$, no significant differences between samples. '2' – $p < 0.005$. The *N* indicates the number of participants for which the corresponding information is available.

Table 2. Cortical region exhibiting greater thinning in MDD group compared to HC

Cluster #	Max -log(p)	Annotation	Size(mm^2)	MNIX	MNIY	MNIZ	CWP
Left hemisphere							
1	-3.389	rostral middle frontal	1599.21	-29.9	54.5	-7.7	0.0002

Max -log(p) indicates the maximum -log(p) value among the vertices in the cluster. CWP = cluster-wise p-value. The Annotation heading refers to the location of the peak voxel as per the Desikan-Killiany atlas. The MNI coordinates indicate the location of the peak vertex.

4.8 Supplementary Material

4.8.1 Participants

Participants were recruited from six clinical sites in Canada: the Djavad Mowafaghian Centre for Brain Health (Vancouver, BC), Hotchkiss Brain Institute (Calgary, AB), University Health Network (Toronto, ON), Centre for Addiction and Mental Health (Toronto, ON), St. Joseph's Healthcare Hamilton (Hamilton, ON) and Providence Care Hospital (Kingston, ON). Recruitment of MDD and HC participants involved referrals from outpatient clinics and advertisements in the community. Subjects included in the MDD group were outpatients aged 18-60 who met DSM-IV-TR criteria for a major depressive episode (MDE) in MDD, the duration of which was longer than 3 months. A confirmation of MDD on the Mini International Neuropsychiatric Interview (MINI) and a score greater than 24 on the Montgomery-Åsberg Depression Scale (MADRS) was used by all clinicians to confirm the MDE. Healthy control subjects were age- and sex-matched and had no history of psychiatric illness. All subjects were fluent in English and free of psychotropic medications for at least 5 half-lives before the baseline visit. Exclusion criteria for MDD subjects were: co-morbid diagnosis of any psychiatric illness considered the primary diagnosis, including bipolar disorder Type I and II or personality disorder that would interfere with participation; high suicidal risk; substance dependence or abuse in past 6 months; neurological or major medical condition; pregnancy or breastfeeding; psychosis in current episode; high risk for hypomanic switch; failure of four or more previous pharmacotherapeutic interventions; previous intolerance

to escitalopram or aripiprazole; having started psychological treatment in past three months with intent to continue and any contraindications to MRI.

4.8.2 MRI acquisition

3.0T structural MRI scans were obtained for each participant at three timepoints: baseline (pre-treatment), week 2 and week 8 from the beginning of treatment. Four different scanner models were used across the 6 clinical sites: Discovery MR750 3.0T (GE Healthcare, Little Chalfont, Buckinghamshire, UK), Signa HDxt 3.0T (GE Healthcare, Little Chalfont, Buckinghamshire, UK), TrioTim 3.0T (Siemens Healthcare, Erlangen, Germany), and Intera 3.0T (Philips Healthcare, Best, Netherlands). A whole-brain T1-weighted turbo gradient echo sequence was used across all sites. The range of specific parameters are as follows: repetition time (TR) = 6.4-1760ms, echo time (TE) = 2.2-3.4s, flip angle = 8-15 degrees, inversion time (TI) = 450-950ms, field of view (FOV) = 220-256mm, acquisition matrix = 256x256 – 512x512, 176-192 contiguous slices at 1mm thickness with voxel dimensions of 1mm^3. Total acquisition time, depending on scanner, ranged from 3:30min to 9:50 minutes. Shortly after scanning, all raw MRI data underwent a first-pass general quality assurance procedure to manually check for artifacts. Scans were accepted or rejected on visual inspection and efforts made to re-scan the participant as permitted by study timeline. Specific details of the multi-site quality control protocol have been described elsewhere [1].

4.8.3 Site correction using ComBat

There are a variety of factors that contribute to site/scanner-specific variations in the overall sample, including different vendors and variations in pulse sequences, image scaling and bias fields. This non-biological variation is the main tradeoff in obtaining enough datapoints for a sufficiently powered neuroimaging analysis and if not well-accounted for, can confound neuroimaging findings [2].

ComBat harmonization involves using linear modelling and empirical Bayes to estimate the additive and multiplicative effects of site on cortical thickness measurements and corrects for them while taking into account any variability associated with biology (sex, age, diagnostic status, etc.) [3,4]. Previous applications of ComBat in cortical thickness studies corrected for site effects in a region-based manner and was found to be reliable in this regard [5]. We will describe here in detail the steps taken to apply ComBat on a vertex-wise basis to accommodate our exploratory whole-brain analyses.

For each subject, FreeSurfer thickness surface maps were resampled to fsaverage, the FreeSurfer standard template, in order to obtain the same number of vertices (features) for each subject and to align all features from all subjects into the same space. Thickness files in fsaverage space were then converted to ASCII format to facilitate loading into Python (version 3.0) using *mris_convert*. We used the Python implementation of ComBat, neuroCombat, for ease of use and increased processing speed, considering the number of features per subject (160 000+). We found that several features (thickness values at certain vertices) were set to 0 for all participants, which precludes proper functioning of the algorithm. Prior to applying the algorithm, we removed these columns from the array, storing the indices for replacement following harmonization. We entered age, age^2 and

sex as covariates and defined either diagnostic group or response status for the *discrete_cols* argument. Data were separately corrected for site effects for each analysis involving different group comparisons to preserve as much of the true biological variability as possible. The harmonized ASCII files were then converted back into FreeSurfer surface file format using the FreeSurfer Matlab scripts *read_ascii_curv.m* and *write_curv.m*. Thickness surface maps were resampled back into native space and smoothing was performed on these files as per usual for vertex-wise GLM analyses.

Figure S1 displays a visual representation of the effect of harmonization in the left rostral middle frontal gyrus, the region of interest identified in our whole-brain analyses. Visual inspection reveals that between-site variation has been reduced while conserving the pattern of between-group differences.

Non-harmonized

Harmonized
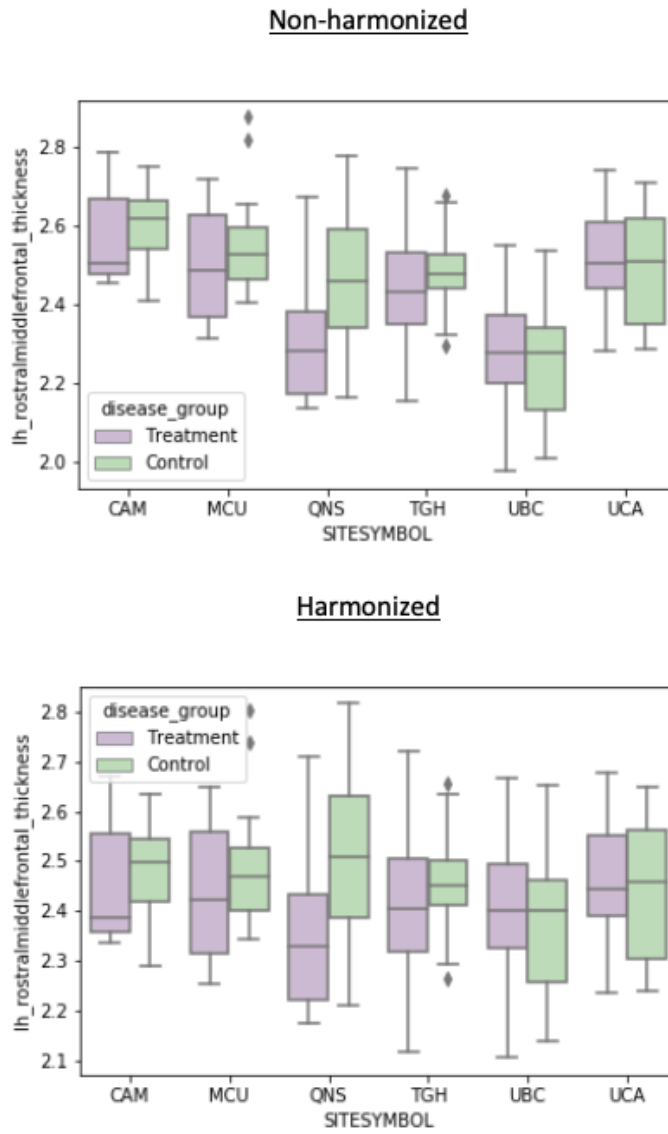
Supplemental Figure 1. Boxplots displaying the distribution of average cortical thickness values in the left rostral middle frontal cortex, categorized by site and stratified by diagnostic group, before and after applying the ComBat harmonization procedure.

Supplemental Table 1. Demographic and clinical information for week-16 MDD response subgroups.

	ESC-16			ARI-16			NR-16		
	N	Mean/ %	SD	N	Mean/ %	SD	N	Mean/ %	SD
Age[1]	64	33.64	11.83	42	36.95	11.96	35	36.91	13.92
Sex % F[1]	64	65.63	--	42	64.29	--	35	54	--
Years of education[1]	63	16.78	2.30	42	17.07	1.92	35	16.66	2.13
Age of illness onset[1]	60	18.82	7.19	41	21.95	10.30	34	21.88	13.30
Illness duration (years)[1]	60	14.58	10.90	41	15.22	14.04	34	15.5	13.03
Baseline MADRS[1]	64	29.37	5.55	42	31.14	5.57	35	29.17	5.02
8-week MADRS[2]	64	7.83	5.15	42	21.88	5.62	35	25.57	8.88
16-week MADRS[2]	64	4.78	4.01	42	8.88	5.20	35	21.8	6.80
Number of previous MDEs[1]	63	3.46	3.35	42	3.07	3.67	33	2.61	3.43

MADRS – Montogomery-Asberg Depression Rating Scale; MDE – major depressive episode; [1] – $p > 0.05$, no significant differences between samples. [2] – $p < 0.005$ in omnibus comparison among the three groups.

Chapter 5: Hypothalamus volume and DNA methylation of stress axis genes in major depressive disorder: A CAN-BIND study report

Jee Su Suh[a,b], Laura M. Fiori[c], Mohammad Ali[a,b], Kate L. Harkness[d], Milita Ramonas[e], Luciano Minuzzi[a,b,f], Stefanie Hassel[g], Stephen C. Strother[h], Mojdeh Zamyadi[h], Stephen R. Arnott[h], Faranak Farzan[j], Jane A. Foster[f,], Raymond W. Lam[k], Glenda M. MacQueen[g], Roumen Milev[l], Daniel J. Müller[m,n], Sagar V. Parikh[p], Susan Rotzinger[m], Roberto B. Sassi[f], Claudio N. Soares[l], Rudolf Uher[q], Sidney H. Kennedy[m,o], Gustavo Turecki[c], Benicio N. Frey[a,b,f]

[a] Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada
[b] Mood Disorders Program and Women's Health Concerns Clinic, St. Joseph's Healthcare Hamilton, ON, Canada
[c] Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Montreal, Canada
[d] Department of Psychology, Queen's University, Kingston, ON, Canada
[e] Juravinski Hospital and Cancer Centre, Hamilton Health Sciences, Hamilton, ON, Canada
[f] Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada
[g] Department of Psychiatry, University of Calgary, Calgary, AB, Canada
[h] Rotman Research Institute, Baycrest, Toronto, ON, Canada
[j] eBrain Lab, School of Mechatronic Systems Engineering, Simon Fraser University, Surrey, British Columbia, Canada
[k] Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada
[l] Departments of Psychiatry and Psychology, Queen's University, and Providence Care Hospital, Kingston, ON, Canada
[m] Department of Psychiatry, University of Toronto
[n] Centre for Addiction and Mental Health, Pharmacogenetics Research Clinic, Toronto, ON, Canada
[o] Centre for Depression and Suicide Studies, St. Michael's Hospital, Toronto ON, Canada
[p] University of Michigan Depression Center, Ann Arbor, MI, United States
[q] Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia

Abstract

Dysfunction of the hypothalamic-pituitary-adrenal (HPA) axis is considered one of the mechanisms underlying the development of major depressive disorder (MDD), but the exact nature of this dysfunction is unknown. We investigated the relationship between hypothalamus volume (HV) and blood-derived DNA methylation in MDD. We obtained brain MRI, clinical and molecular data from 181 unmedicated MDD and 90 healthy control (HC) participants. MDD participants received a 16-week standardized antidepressant treatment protocol, as part of the first Canadian Biomarker Integration Network in Depression (CAN-BIND) study. We collected bilateral HV measures via manual segmentation by two independent raters. DNA methylation and RNA sequencing were performed for three key HPA axis-regulating genes coding for the corticotropin-binding protein (CRHBP), glucocorticoid receptor (NR3C1) and FK506 binding protein 5 (FKBP5). We used elastic net regression to perform variable selection and assess predictive ability of methylation variables on HV. Left HV was negatively associated with duration of current episode (ρ=-0.17, p=0.035). We did not observe significant differences in HV between MDD and HC or any associations between HV and treatment response at weeks 8 or 16, overall depression severity, illness duration or childhood maltreatment. We also did not observe any differentially methylated CpG sites between MDD and HC groups. After assessing functionality by correlating methylation levels with RNA expression of the respective genes, we observed that the number of functionally relevant CpG sites differed between MDD and HC groups in FKBP5 ($\chi^2 = 77.25$, p<0.0001) and NR3C1 ($\chi^2 = 7.29$, p=0.007). Cross-referencing functionally relevant CpG

sites to those that were highly ranked in predicting HV in elastic net modelling identified

one site from FKBP5 (cg03591753) and one from NR3C1 (cg20728768) within the MDD

group. Stronger associations between DNA methylation, gene expression and HV in

MDD suggest a novel putative molecular pathway of stress-related sensitivity in

depression. Future studies should consider utilizing the epigenome and ultra-high field

MR data which would allow the investigation of HV sub-fields.

5.1 Introduction

The ongoing search for biomarkers and predictors of response in major depressive disorder (MDD) has seen varying levels of success from separate clinical, molecular and neuroimaging perspectives (Kang and Cho 2020; Perna et al. 2020; Fiori et al. 2018; Mora et al. 2018). In the present study, we combined and examined relationships between a priori-defined clinical, neuroimaging and molecular variables within the theoretical framework of the hypothalamic-pituitary-adrenal (HPA) stress axis, posited to be dysregulated in MDD (Cernackova et al. 2020; Halaris 2019).

The HPA axis is comprised of a series of hormonal actions and feedback regulation steps primarily among the hypothalamus, pituitary and adrenal glands. Corticotropin-releasing hormone (CRH) is excreted by the hypothalamus into portal blood, stimulating the anterior pituitary to release adrenocorticotropic hormone (ACTH) that results in cortisol release from the adrenals. Cortisol exerts negative feedback on this axis by binding to glucocorticoid receptors (GR) located in the hypothalamus and pituitary, as well as other central regions such as the hippocampus and prefrontal cortex. A plausible etiological theory of MDD is the desensitization of GR to cortisol due to inflammation, resulting in a dysregulated and prolonged stress response (Pace, Hu, and Miller 2007). FK506 binding protein 51 (FKBP5) is a chaperone to GR, and the two have been implicated in multiple studies with respect to MDD susceptibility and risk (Roy, Shelton, and Dwivedi 2017; Rao et al. 2016; Piechaczek et al. 2019), antidepressant response (Ising et al. 2019; Binder et al. 2004) and childhood maltreatment (Bockmühl et al. 2015). Although some studies have examined the interaction of molecular variables

pertaining to these genes with structural and functional neuroimaging (Tozzi et al. 2018, for example), none have done so in relation to hypothalamus volumetry (HV). Of particular note is the lack of studies examining the link between HV and childhood maltreatment in the imaging literature, although recent preclinical work suggests that early life stress suppresses cellular proliferation in the hypothalamus in adulthood (Bielefeld et al. 2021).

There has been some success identifying volumetric indicators of MDD symptoms and antidepressant response in vivo (Kang and Cho 2020), established most reliably in the hippocampus (Nogovitsyn et al. 2019; Santos et al. 2018). Additionally, antidepressant treatment has been associated with reversal of hippocampal volume deficits to a certain extent on a long-term basis (Malberg, Hen, and Madsen 2021; Boldrini et al. 2013). Both the hippocampus and the hypothalamus express significant levels of GR, and although it is widely known that the hippocampus exhibits neurogenesis, the discovery of the same phenomenon in the hypothalamus is relatively more recent and less explored (Markakis et al. 2004; Cheng 2013). Despite this and its central role in stress regulation, the hypothalamus is not studied often volumetrically in MDD due to its amorphous, heterogeneous structure leading to difficulty in segmentation. A few studies have examined group differences in volume in both in vivo and post-mortem contexts that were discrepant due to methodological limitations (Schindler et al. 2012). Post-mortem studies (using the same patient sample) found bilaterally smaller volumes of the hypothalamus in MDD (Bielau et al. 2005; Bernstein et al. 2012). Similar results were found in studies examining hypothalamus volume in anxiety (Terlevic et al. 2013) and

schizophrenia (Koolschijn et al. 2008). However, a recent imaging study found greater left-sided volume of this region in MDD in a small sample of 20 MDD participants (Schindler et al. 2019). We aim to leverage a larger and well-defined MDD study sample to add to these results.

The objective of this study was to investigate the relationship between HPA-axis-related brain and blood variables in MDD in an exploratory fashion. We used bilateral HV and the blood DNA methylation and gene expression profiles of three genes involved in the HPA axis: CRH binding protein (CRHBP) as a proxy for CRH, NR3C1 (encodes GR) and FKBP5. We used the elastic net regularization method to quantify the extent to which DNA methylation at individual CpG sites on these genes predict HV. Secondary objectives included a hypothesis-driven analysis for group differences in HV measures and their clinical correlations (including childhood maltreatment, which has not yet been studied in relation to HV) in the largest analysis on this brain structure to date, for which we predicted smaller volumes in MDD. An additional discovery analysis was performed to assess whether there are any longitudinal changes in HV over 8 weeks of escitalopram treatment.

5.2 Methods

5.2.1 Study protocol

Participant recruitment occurred at six Canadian research sites from August 2013 to December 2016. All protocols were approved by the Research Ethics Boards at each institution. MDD participants were unmedicated at baseline, aged 18-60 years and met DSM-IV-TR criteria for a major depressive episode. MDD participants received flexible-

dosage escitalopram treatment for 8 weeks, starting at 10mg with a maximum dose of 20mg per day. Clinical response was defined as a greater than 50% reduction in the Montgomery-Asberg Depression Rating Scale (MADRS) score after 8 weeks. At the 8-week timepoint, those who did not achieve response received an adjunctive treatment of aripiprazole (flexible dosage 2-10mg/day); responders continued escitalopram treatment as originally prescribed. At 16 weeks, MDD participants were further categorized into the following response categories: early sustained responders (achieved response at week 8, continued to respond at week 16), late responders (did not achieve week 8 response, but achieved response at week 16 with adjunctive treatment) and non-responders (week 8 non-responders that failed to respond to adjunctive treatment by week 16). Additional inclusion/exclusion criteria for the CAN-BIND 1 study can be found in Supplementary Materials. Further details pertaining to clinical, neuroimaging and biomarker protocols as well as patient outcomes and a CONSORT diagram outlining the flow of participants throughout the 16-week trial are available elsewhere (Kennedy et al. 2019; Lam et al. 2016; MacQueen et al. 2019).

5.2.2 Image acquisition

Structural T1-weighted images were obtained at 3T at baseline and week 8. Across the six clinical sites four scanner types were used: three GE Discovery MR750, GE Signa HDxt, Siemens TrioTim and Philips Intera. All sites used a whole-brain turbo gradient echo sequence with the following ranges of parameters: repetition time (TR) = 6.4-1760ms, echo time (TE) = 2.2-3.4s, flip angle = 8-15 degrees, inversion time (TI) =

450-950ms, field of view (FOV) = 220-256mm, acquisition matrix = 256x256 – 512x512, 176-192 contiguous slices at 1mm thickness with voxel dimensions of 1mm isotropic. Total acquisition time ranged from 3:30min to 9:50 minutes. Raw data were manually checked for artifacts and if the visual inspection warranted, efforts were made to re-scan the participant as permitted by study timeline. Refer to MacQueen et al. (2019) for specific details on the neuroimaging protocols.

5.2.3 Manual hypothalamus segmentation

The retrospective nature of our investigation of this brain region and the variability in our 3T MRI data due to multiple scanners required manual segmentation to ensure data quality. We drew upon many recent sources (described in detail below) pertaining to hypothalamic boundaries and strengthened our protocol by blinding raters to subject identity/attributes.

Manual segmentation of the hypothalamus was performed by two independent raters (J.S. & M.A.) using the open-source software ITK-SNAP (http://www.itksnap.org/). Segmentation was aided by an overlay of the subject's gray matter tissue probability map (GM-TPM) as per Wolff et al. (2018), automatically generated using the VBM-Segment tool in SPM 12 in Matlab (with very light regularization, all other default settings maintained). Raters were blinded to subject ID, scanning site, treatment group and scanning timepoint. Subjects were shuffled and images were segmented in random order. A training phase for raters comprised 100+

'throwaway' segmentations prior to proper data collection of segmentation values to overcome any practice effects.

The segmentation boundaries from Bocchetta et al. (2015) provided a visual guide for manual work, where the authors used co-registration of T1 and T2-weighted images to discern finer details and subdivide the segmentation into subunits. To make up for our lack of T2-weighted images, for each subject we used the GM-TPM as a transparent overlay (opacity ~6-10%) for additional perspective (Figure 1). We deviated slightly from the protocol of Bocchetta et al. (2015). While that group sub-segmented and subsequently excluded the fornix from volume calculations, we included the fornix in our segmentation once the medial edge separated from the wall of the third ventricle, thereby aligning more closely with the inclusion criteria in Schindler et al. (2013). The reason for this deviation is that in our 3T dataset, the fornix becomes difficult to distinguish from the surrounding gray matter as it starts to travel away from the ventricle wall in more caudal slices. Detailed anatomical landmarks are tabulated in table e1 of the supplement in Bocchetta et al. (2015).

We carried out voxel-wise segmentation in itk-SNAP (brush size 1, isotropic) to 'shade in' the right and left hemispheres of the hypothalamus separately, working in the coronal view in the rostral-caudal direction with sagittal and horizontal views visible for reference. We alternated starting with the left or right hemispheres and for each image and hid the first side from view as the other hemisphere was segmented to avoid visual bias. Whole volume segmentations included the preoptic region (starting at the anterior commissure), anterior hypothalamic region and mammillary bodies (stopping at the last

'complete' slice). We randomly selected 10-15% of segmentations from each site as well as a select number of more challenging segmentations (N= 50 segmentations in total) for inspection by a neuroanatomist/radiologist (M.R.), who provided expert quality assurance feedback on the boundaries.

Inter-rater agreement scores were calculated across all segmentations and intra-rater agreement was calculated independently for each rater using a randomly selected set of 50 segmentations. Spatial overlap was calculated using Dice scores and numeric agreement was calculated using two-way mixed, single score intraclass correlation coefficient (ICC; Shrout and Fleiss convention [3,1]). The average of the values from each independent rater were used in subsequent analyses. Volumes were corrected for intracranial volume (ICV) using the power-proportion method as described in Liu et al. (2014), shown to exhibit better performance in removing the confound of ICV compared to scaling by a linear factor only. The scaling factor s was defined as the beta coefficient for log(HV) regressed on log(ICV) for each hemisphere. The resulting ICV-corrected HV values were obtained by dividing the original values by ICV^s. A measure of hemispheric asymmetry was calculated as (Left volume - Right volume)/Total volume (Zuo et al. 2019).

5.2.4 Molecular analyses of RNA transcript levels and DNA methylation

For all participants, molecular analyses were performed at McGill University and the Genome Quebec Innovation Centre; full methodological details pertaining to DNA methylation analyses can be found in Ju et al. (2019). Briefly, genome-wide DNA

methylation was performed on the Infinium MethylationEPIC Beadschip (Illumina, USA) using DNA extracted from whole blood samples obtained from participants at baseline prior to treatment. Quality assurance and pre-processing were performed using the Chip Analysis Methylation Pipeline Bioconductor package in R 3.4. Beta values at each CpG site refer to the ratio of methylated signal to the sum of unmethylated and methylated signal.

For gene expression analyses, whole blood for RNA was collected in EDTA tubes and filtered using LeukoLOCK filters (Life Technologies, USA). Total RNA was extracted using a modified version of the LeukoLOCK Total RNA Isolation System protocol and included DNase treatment to remove genomic DNA. RNA quality was assessed using the Agilent 2200 Tapestation, and only samples with RNA Integrity Number (RIN) ≥ 6.0 were used. All libraries were prepared using the Illumina TruSeq mRNA stranded protocol following the manufacturer's instructions. Samples were sequenced at McGill University and the Genome Quebec Innovation Centre (Montreal, Canada) using the Illumina HiSeq4000 with 100nt paired-end reads. FASTXToolkit and Trimmomatic were respectively used for quality and adapter trimming. Tophat2, using bowtie2 was used to align the cleaned reads to reference genome. Reads that lost their mates through the cleaning process were aligned independently from the reads that still had pairs. Quantification on each gene's expression was estimated using HTSeq-count and a reference transcript annotation from ENSEMBL. Counts for the paired and orphaned reads for each sample were added to each other. Normalization was conducted on the resulting gene matrix using DESeq2.

5.2.5 Childhood maltreatment measure

The Childhood Experience of Care and Abuse (CECA) interview was used to assess childhood maltreatment and caregiver relationships up to age 18 (Bifulco, Brown, and Harris 1994; Chakrabarty et al. 2019). CECA is retrospective and semi-structured, administered via secure videoconferencing and rated using standardized criteria. Subscales of antipathy, neglect, physical abuse and sexual abuse were rated on a 4-point severity scale based on narrative details such as frequency, age of experiencing the maltreatment and for how long, relationship with perpetrator and degree of injury. Subscales were summed to yield an overall childhood maltreatment score.

5.2.6 Statistical analyses

All statistical analyses were performed in Python 3.8.1 using packages *statsmodels*, *sklearn and pingouin* (Vallat, 2018). We tested for significant differences between demographic variables using either Student's t-test or chi-square test for continuous and categorical variables, respectively. Testing for group differences in HV measures (left [LHV], right [RHV], asymmetry) was performed using Student's t-test (in order to report confidence intervals and effect sizes) and subsequently general linear model to further control for sex and age. Spearman's rank correlations with each HV measure, which were normally distributed, were performed with the following clinical measures, which were not normally distributed: baseline symptom severity (MADRS; individual items and total score), childhood maltreatment (CECA; sub-scores and total score) current episode duration, age of MDD onset and number of prior episodes.

Baseline HV variables were compared between antidepressant response groups at week 8 (responders, non-responders) and week 16 (early sustained responders, late responders, non-responders) to determine whether baseline values are associated with treatment response. Linear mixed models were used for the longitudinal analysis of HV change over 8 weeks, during which the MDD group received escitalopram treatment, accounting for sex and age. Site effects for both HV and methylation were assessed with analysis of variance. Correlations with age and sex differences were exploratory tests performed for right, left and asymmetry HV measures.

We assessed variances for each CpG site, excluding sites whose standard deviation fell below the threshold of 0.02. All analyses described hereon use only CpG sites that passed this variance threshold. After checking homoscedasticity using Bartlett's test, Student's t-tests were used to compare methylation levels between MDD and HC groups at each CpG site, corrected for multiple comparisons using false discovery rate (FDR) at p=0.05. To assess functional significance of CpG sites for each gene, we computed Spearman's ρ on the normalized beta values of each CpG site and the expression level of their respective genes, within each group. Statistical significance of correlations for the MDD group, which comprised twice as many samples as the HC group, were confirmed with 95% confidence intervals of ρ with a bootstrapping procedure that used a resampling size identical to that of the HC group. Correlations were further corrected for multiple comparisons using false discovery rate (FDR) at p=0.05. Chi-square test was used to assess the significance of the group difference in the numbers of significant gene expression/methylation correlations.

To explore relationships between HV measures and methylation profiles (low variance sites excluded), we used multiple linear regression with elastic net regularization (Zou and Hastie 2005). Regularization methods optimize model coefficients such that the most important variables are assigned higher values. Due to combining variable selection and parameter estimation in one step, these methods do not run the risk of inflated standard errors and model variances as traditional ordinary least squares methods do (Finch and Hernandez Finch 2016) and are additionally well-suited for the common scenario where the number of predictors is smaller than or roughly equivalent to the number of samples.

Models were constructed for left and right HV each as a dependent variable and in each diagnostic group separately, resulting in 4 elastic net models (MDD–LHV, MDD–RHV, HC–LHV, HC–RHV). Sex and age were included in addition to methylation profiles at each CpG site, comprising 64 explanatory variables in total. For each model, we split the data into train/test samples using a 0.75:0.25 ratio and used 10-fold cross-validation and grid search for hyperparameter optimization. We repeated this procedure 100 times for each model and recorded values for mean squared error (MSE), mean absolute error (MAE) and R^2 on the test sample at each iteration. To calculate relative variable importance (RVI) scores, we counted how many times out of 100 runs the absolute value of each variable's coefficient ranked in the top 15% of model coefficients. We cross-referenced CpG sites with RVI scores > 50 with those that were significantly correlated with gene expression (indicating functional significance) in order to infer which sites might reasonably be associated with HV. Although we are not able to directly

infer p-values for coefficients, the procedure of regularization in itself is designed to shrink 'non-significant' coefficients toward zero; additionally, we included sex and age as reference variables to infer relative strength of the other methylation variables in predicting HV.

5.3 Results

274 images from baseline and 223 images from week 8 were ultimately included for final analyses after excluding for visual quality issues (e.g., failing GM-TPM generation, overly pixelated quality) for a total of 497 manual segmentations. The outlier analysis resulted in removal of 3 participants for a total of 271 participants (181 MDD, 90 HC) at baseline and 217 participants at week 8 (138 MDD, 79 HC, no outliers found). There were no significant differences between groups in mean age, female:male ratio or ethnicity. Groups significantly differed in years of education, baseline MADRS, % family history of psychiatric disorders and total CECA score (Table 1).

5.3.1 Demographics

A one-way ANOVA revealed no effect of site on methylation values but did indicate an effect on HV values. Further pairwise testing revealed that one site (MCU) exhibited significantly different HV values in relation to all other sites. No significant differences were observed among the other 5 sites. A statistical correction method, ComBat (Fortin et al. 2017), was used to eliminate the effect of site while retaining variance associated with diagnosis, sex and age (Supplementary Figure 1); however,

using site-corrected values did not affect statistical significance of results. As the estimation of site effects is inherently noisy and there is no guarantee that we corrected for only the variance truly associated with scanning site, we report uncorrected values to prioritize simpler modelling (Kass et al. 2016). Refer to Supplemental Table 1 for demographic and clinical information grouped by site.

5.3.2 Inter- and intra-rater agreement

Raters exhibited good agreement with themselves and between each other overall with values equal to or greater than 0.78. Refer to Supplemental Table 2 for inter- and intra-rater agreement statistics.

5.3.3 Demographic and clinical associations with HV

GLM analyses (satisfying statistical assumptions of normality, linearity and homoscedasticity) accounting for sex and age revealed no significant difference between MDD and HC in either RHV or LHV (Figure 2A; see Supplemental Table 3 for t-test results, power analyses and effect sizes for mean differences). There was a negative association of LHV with duration of current episode (Spearman's ρ=-0.166, p=0.03). With respect to individual MADRS items, we observed for LHV a significant positive correlation with *difficulty concentrating* (ρ=0.16, p=0.035). For RHV, the negative correlation with *inner tension* was significant (ρ=-0.19, p=0.011). However, these clinical associations did not survive FDR correction for comparisons made across all items. There was no association of HV variables with CECA scores, baseline total MADRS, illness

duration, family history of psychiatric disorders or antidepressant response status at 8 or 16 weeks, accounting for age and sex.

Within the MDD group, a significant difference between males and females was observed in RHV (t=2.13, p=0.044) but not LHV (t=1.64, p=0.10). This pattern was similar in the HC group with no sex difference observed in LHV (t=1.81, p=0.08) although the difference in RHV was marginally significant (t=1.96, p=0.05) (Figure 2B). Significant associations with age were observed in both left and right HV in the MDD group (LHV: ρ=-0.27, p=2.9e-4; RHV: ρ=-0.24, p=1.3e-3) but not in the HC group. Slopes were similar in controls but not statistically significant (likely due to an uneven age distribution at higher values in this group). There was no effect of group on the association with age in either hemisphere (Figure 2C).

Longitudinal analysis revealed no significant changes in HV variables over 8 weeks of escitalopram treatment in the MDD group. Visualization of baseline and week 8 values reveal a small yet consistent group difference that persisted after 8 weeks (Supplementary Figure 2), but these differences were not statistically significant at either timepoint.

5.3.4 Differentially methylated points between groups

3 CpG sites exhibited reduced methylation in MDD (cg19645279 on *NR3C1*: t=2.95, p=0.003; cg18718518 on *NR3C1*: t=2.56, p=0.011; cg07633853 on *FKBP5*: t=2.46, p=0.015), although none of these comparisons survived FDR correction.

5.3.5 Functional relevance analysis of CpG sites

We excluded 15 out of 23 CpG sites on *CRHBP*, 24 out of 44 sites on *FKBP5* and

54 out of 81 on *NR3C1* from analyses due to low variance (SD<0.02). Within the MDD

group, 11 out of the remaining 20 CpG sites on *FKBP5* correlated significantly to gene

expression levels, and 8 out of 27 were significantly correlated for *NR3C1*. Within the HC

group, the methylation profiles of 4 CpG sites on *NR3C1* were significantly correlated

with gene expression (Figure 3; Supplemental Table 4). Chi-square tests confirmed that

the number of functionally relevant CpG sites within *NR3C1* and *FKBP5* in the MDD

group were significantly different from the number of significant correlations observed in

HC (*FKBP5*: χ^2 = 77.25, p<0.0001; *NR3C1*: χ^2 = 7.29, p=0.007). Neither group exhibited

any functionally relevant CpG sites within *CRHBP*.

5.3.6 Elastic net modelling of the effect of DNA methylation on HV

On average across runs, regularized models for predicting RHV and LHV in

MDD explained significantly higher amounts of variance (R^2) than in HC (LHV: Welch's

t=10.91, p=2.63e-20; RHV: Welch's t=8.36, p=2.3e-14) (Table 2; Figure 4). The variable

of sex was shown to be consistently highly ranked across all four models, although age

ranked highly only in the HC models. In the MDD group only, two highly ranked CpG

sites were shown to also be significantly correlated with gene expression (cg20728768

[*NR3C1*], cg03591753 [*FKBP5*]). The coefficient signs for both variables indicated

positive associations of methylation with HV. *CRHBP* featured most heavily among

highly ranked variables in the MDD models (4 CpG sites) with the other 4 sites evenly

split among *NR3C1* and *FKBP5*, whereas all highly ranked CpG sites in the HC models were located on *FKBP5*.

5.4 Discussion

We investigated characteristics of hypothalamus volume and epigenetic/gene expression profiles of three specific genes with major roles in the HPA axis in MDD. Novel findings of this study include the identification of two functionally relevant CpG sites that were predictive of HV within the MDD group, but not in healthy individuals, as well as the negative association of LHV with current episode duration. We showed that there are distinct relationships between HV and sex/age that are not affected by MDD status. We did not find any indication of a change in HV associated with 8-week escitalopram treatment. Finally, we did not observe that baseline volume was predictive of antidepressant response status at 8 or 16 weeks.

We found distinct differences between groups on the correlations between methylation and gene expression, whereby MDD exhibited more CpG sites within each gene that significantly correlated with its RNA expression levels; however, reduced methylation levels at 3 CpG sites (2 located on *NR3C1* and 1 on *FKBP5*) in MDD did not survive correction for multiple comparisons. Elastic net models further quantified the relative contributions of each CpG site in predicting HV. One of the highly ranked CpG sites that was also significantly correlated with gene expression was cg20728768, located on NR3C1. Previously, methylation at this site was found to be positively correlated with a measure of resilience in a sample with varying levels of PTSD symptom severity

(Miller et al. 2020). In our study, methylation at this site was positively associated with HV, consistent with the implication that it may be a protective factor. In another study, it was negatively associated with PTSD, although this correlation did not survive correction for multiple comparisons (Mehta et al. 2019). The other significant CpG site was identified to be cg03591753, located on FKBP5. In one previous study, methylation at this site significantly predicted cortisol reactivity in interaction with the rs1360780 variant of the FKBP5 gene and resistant behaviour to predict cortisol reactivity in infants (Mulder et al. 2017). We add to these findings, positing that DNA methylation levels at these CpG sites are linked to hypothalamic macrostructure in depression.

The same explanatory variables, including sex and age as reference variables, explained significantly less model variance overall within the HC group. The HC models also consistently ranked sex and age highly among explanatory variables in predicting HV, although age was not significantly correlated with HV in HC as indicated by the Spearman correlation; this gives some indication as to the relative strength of the association of HV with methylation variables that were ranked below age in HC compared to MDD (ie. relatively weaker). Conversely, although age was significantly correlated with HV in the MDD group, its RVI was <20 in both MDD elastic net models, indicating that the methylation variables that ranked higher were relatively more successful in predicting RHV and LHV than age. Additionally, only the MDD models identified highly ranked sites that were also significantly correlated with gene expression. Taken together, these results demonstrate greater overall statistical associations among variables of gene expression, DNA methylation and hypothalamic structure in MDD than

in HC. This may be indicative of a concerted physiological response in MDD that could either be 1) a developmental or environment-elicited disruption leading to MDD symptoms or 2) a compensatory mechanism by the body to counteract other underlying etiological factors. With only baseline data available for DNA methylation, we were not able to test causal hypotheses. Although we did not observe group differences in DNA methylation among the CpG sites in these 3 genes, this was not entirely surprising given other genes have been identified to exhibit differential methylation in MDD (Chan et al. 2020; Xie et al. 2021).

Not observing significant group differences in either left nor right HV suggests that either there are no hypothalamic volumetric changes in MDD, or the volumetric changes might be very small and would require much larger samples to be detected. The latter is likely the case, as indicated by power analyses (Supplementary Table 3). In the only other study that explicitly examined HV in MDD *in vivo*, Schindler et al. (2019) found essentially the opposite result: their sample of unmedicated MDD participants exhibited greater left HV in a much smaller cohort of only 20 MDD and 20 HC participants. Other relevant major differences between our analyses aside from sample size are the correction for ICV using the power-proportion method and including age as a covariate given that we observed a significant HV-age correlation in the MDD sample. Otherwise, the lack of association between HV and other clinical variables is consistent across the two studies. We additionally did not find an association between HV and CECA, the composite score of childhood maltreatment. As there are no previous studies examining this relationship available for direct comparison, we entertain a few

explanations: 1) early childhood experiences of emotional, physical or sexual abuse are not reflected by hypothalamic structure, 2) the CECA score, with its 4-point sub-scores, does not capture the specific lasting aspects of trauma that may be embedded in HV, or 3) similar to depressive status, more power is required to detect a finer-grained relationship between early trauma and HV.

Previous literature regarding the link between HPA axis variables and depressive characteristics has been equivocal. It has previously been observed that most depression characteristics (severity, chronicity, symptom profile, prior childhood trauma) were not associated with HPA axis activity except for comorbid anxiety (Vreeburg et al. 2009). A more recent meta-analysis assessing effects across 361 studies did observe that patients with depression exhibited increased cortisol and adrenocorticotropic hormone levels; however, this effect did not extend to CRH, with the additional caveat that methodological limitations may have inflated some effects among primary studies (Stetler and Miller 2011). That we observed no change in volume due to escitalopram treatment is not entirely surprising given the heterogeneous and contradictory evidence for detectable volumetric changes in humans following pharmacological treatment in MDD (Enneking et al. 2020; Suh et al. 2020).

With respect to the imaging component, we adapted innovative methods from previous papers (Wolff et al. 2018; Bocchetta et al. 2015; Schindler et al. 2013) to support a comprehensive and highly structured manual segmentation protocol designed to mitigate the limitations inherent in the study (eg. 3T resolution, scanner variability, lack of T2-weighted images for contrast, inter-rater agreement). A proportion of segmentations

was also checked for boundary accuracy and overall quality by a certified neuroradiologist with extensive neuroanatomical expertise. Given our data were derived from 3T imaging from multiple scanners, we did not have sufficient resolution to do a subunit-level analysis (Bocchetta et al. 2015). In particular, since the paraventricular nucleus is solely responsible for the release of CRH, being able to isolate the volume of this region may have increased our chances of detecting a relationship with GE and/or DNA methylation. A major recommendation for future research is to use higher resolution images coupled with a robust semi-automatic segmentation procedure in order to be able to discern hypothalamic sub-fields. Another common limitation in our study, as in other human *in vivo* studies, is that despite having access to structural and functional data *in vivo*, it was not possible to obtain the direct genetic and proteomic correlates in the human brain. Future studies could also explore the role of pro-inflammatory markers (Hiles et al. 2012; Milenkovic et al. 2019; Gill et al. 2020), notwithstanding the current challenges in integrating multimodal information in biological psychiatry.

5.5 Conclusion

In a large sample of well-characterized individuals with major depression, we investigated baseline hypothalamic volume characteristics in MDD and whether they are correlated to DNA methylation and expression of candidate genes implicated in HPA axis function. We identified CpG sites whose methylation levels simultaneously exhibited high scores in predicting HV and were significantly correlated to gene expression in MDD but not in HC, which may suggest a molecular pathway of stress-related sensitivity in depression. Future studies could expand their genetic approach to either the entire

epigenome or other/additional candidate genes, as well as utilizing increasingly available ultra-high-resolution images at 7T or higher to discern hypothalamic sub-fields.

5.6 Acknowledgements

5.7 Declarations of Interest

5.8 References

Bernstein, Hans-Gert, Melanie Klix, Henrik Dobrowolny, Ralf Brisch, Johann Steiner, Hendrik Bielau, Tomasz Gos, and Bernhard Bogerts. 2012. "A Postmortem Assessment of Mammillary Body Volume, Neuronal Number and Densities, and Fornix Volume in Subjects with Mood Disorders." *European Archives of Psychiatry and Clinical Neuroscience* 262 (8): 637–46. https://doi.org/10.1007/s00406-012-0300-4.

Bielau, Hendrik, Kurt Trübner, Dieter Krell, Marcus W. Agelink, Hans-Gert Bernstein, Renate Stauch, Christian Mawrin, et al. 2005. "Volume Deficits of Subcortical Nuclei in Mood Disorders A Postmortem Study." *European Archives of Psychiatry and Clinical Neuroscience* 255 (6): 401–12. https://doi.org/10.1007/s00406-005-0581-y.

Bielefeld, Pascal, Maralinde R. Abbink, Anna R. Davidson, Niels Reijner, Oihane Abiega, Paul J. Lucassen, Aniko Korosi, and Carlos P. Fitzsimons. 2021. "Early Life Stress Decreases Cell Proliferation and the Number of Putative Adult Neural Stem Cells in the Adult Hypothalamus." *Stress (Amsterdam, Netherlands)* 24 (2): 189–95. https://doi.org/10.1080/10253890.2021.1879787.

Bifulco, A., G. W. Brown, and T. O. Harris. 1994. "Childhood Experience of Care and Abuse (CECA): A Retrospective Interview Measure." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 35 (8): 1419–35.

Binder, Elisabeth B., Daria Salyakina, Peter Lichtner, Gabriele M. Wochnik, Marcus Ising, Benno Pütz, Sergi Papiol, et al. 2004. "Polymorphisms in FKBP5 Are

Associated with Increased Recurrence of Depressive Episodes and Rapid Response to Antidepressant Treatment." *Nature Genetics* 36 (12): 1319–25. https://doi.org/10.1038/ng1479.

Bocchetta, Martina, Elizabeth Gordon, Emily Manning, Josephine Barnes, David M. Cash, Miklos Espak, David L. Thomas, et al. 2015a. "Detailed Volumetric Analysis of the Hypothalamus in Behavioral Variant Frontotemporal Dementia." *Journal of Neurology* 262 (12): 2635–42. https://doi.org/10.1007/s00415-015-7885-2.

Bockmühl, Yvonne, Alexandre V. Patchev, Arleta Madejska, Anke Hoffmann, Joao C. Sousa, Nuno Sousa, Florian Holsboer, Osborne F. X. Almeida, and Dietmar Spengler. 2015. "Methylation at the CpG Island Shore Region Upregulates Nr3c1 Promoter Activity after Early-Life Stress." *Epigenetics* 10 (3): 247–57. https://doi.org/10.1080/15592294.2015.1017199.

Boldrini, Maura, Adrienne N. Santiago, René Hen, Andrew J. Dwork, Gorazd B. Rosoklija, Hadassah Tamir, Victoria Arango, and J. John Mann. 2013. "Hippocampal Granule Neuron Number and Dentate Gyrus Volume in Antidepressant-Treated and Untreated Major Depression." *Neuropsychopharmacology* 38 (6): 1068–77. https://doi.org/10.1038/npp.2013.5.

Cernackova, Alena, Zdenka Durackova, Jana Trebaticka, and Boris Mravec. 2020. "Neuroinflammation and Depressive Disorder: The Role of the Hypothalamus." *Journal of Clinical Neuroscience* 75 (May): 5–10. https://doi.org/10.1016/j.jocn.2020.03.005.

Chakrabarty, Trisha, Kate L. Harkness, Shane J. McInerney, Lena C. Quilty, Roumen V. Milev, Sidney H. Kennedy, Benicio N. Frey, et al. 2019. "Childhood Maltreatment and Cognitive Functioning in Patients with Major Depressive Disorder: A CAN-BIND-1 Report." *Psychological Medicine*, October, 1–12. https://doi.org/10.1017/S003329171900268X.

Chan, Robin F., Gustavo Turecki, Andrey A. Shabalin, Jerry Guintivano, Min Zhao, Lin Y. Xie, Gerard van Grootheest, et al. 2020. "Cell Type–Specific Methylome-Wide Association Studies Implicate Neurotrophin and Innate Immune Signaling in Major Depressive Disorder." *Biological Psychiatry*, Mechanisms of Major Depression, 87 (5): 431–42. https://doi.org/10.1016/j.biopsych.2019.10.014.

Cheng, Mei-Fang. 2013. "Hypothalamic Neurogenesis in the Adult Brain." *Frontiers in Neuroendocrinology* 34 (3): 167–78. https://doi.org/10.1016/j.yfrne.2013.05.001.

Enneking, Verena, Elisabeth J. Leehr, Udo Dannlowski, and Ronny Redlich. 2020. "Brain Structural Effects of Treatments for Depression and Biomarkers of Response: A Systematic Review of Neuroimaging Studies." *Psychological Medicine* 50 (2): 187–209. https://doi.org/10.1017/S0033291719003660.

Finch, W. Holmes, and Maria E. Hernandez Finch. 2016. "Regularization Methods for Fitting Linear Models with Small Sample Sizes: Fitting the Lasso Estimator Using R." *Practical Assessment, Research and Evaluation* 21 (7). https://doi.org/10.7275/JR3D-CQ04.

Fiori, Laura M., Rixing Lin, Chelsey Ju, Raoul Belzeaux, and Gustavo Turecki. 2018. "Using Epigenetic Tools to Investigate Antidepressant Response." *Progress in

Molecular Biology and Translational Science 158: 255–72.
https://doi.org/10.1016/bs.pmbts.2018.04.004.

Fortin, Jean-Philippe, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem
Aselcioglu, Phil Adams, Crystal Cooper, et al. 2017. "Harmonization of Cortical
Thickness Measurements across Scanners and Sites." *BioRxiv*, June.
https://doi.org/10.1101/148502.

Gill, Hartej, Sabine El-Halabi, Amna Majeed, Barjot Gill, Leanna M. W. Lui, Rodrigo B.
Mansur, Orly Lipsitz, et al. 2020. "The Association Between Adverse Childhood
Experiences and Inflammation in Patients with Major Depressive Disorder: A
Systematic Review." *Journal of Affective Disorders* 272 (July): 1–7.
https://doi.org/10.1016/j.jad.2020.03.145.

Halaris, Angelos. 2019. "Inflammation and Depression but Where Does the Inflammation
Come From?" *Current Opinion in Psychiatry* 32 (5): 422–28.
https://doi.org/10.1097/YCO.0000000000000531.

Hiles, Sarah A., Amanda L. Baker, Theo de Malmanche, and John Attia. 2012. "A Meta-
Analysis of Differences in IL-6 and IL-10 between People with and without
Depression: Exploring the Causes of Heterogeneity." *Brain, Behavior, and
Immunity* 26 (7): 1180–88. https://doi.org/10.1016/j.bbi.2012.06.001.

Ising, Marcus, Giuseppina Maccarrone, Tanja Brückl, Sandra Scheuer, Johannes
Hennings, Florian Holsboer, Christoph W. Turck, Manfred Uhr, and Susanne
Lucae. 2019. "FKBP5 Gene Expression Predicts Antidepressant Treatment

Outcome in Depression." *International Journal of Molecular Sciences* 20 (3). https://doi.org/10.3390/ijms20030485.

Ju, Chelsey, Laura M. Fiori, Raoul Belzeaux, Jean-Francois Theroux, Gary Gang Chen, Zahia Aouabed, Pierre Blier, et al. 2019. "Integrated Genome-Wide Methylation and Expression Analyses Reveal Functional Predictors of Response to Antidepressants." *Translational Psychiatry* 9 (1): 254. https://doi.org/10.1038/s41398-019-0589-0.

Kang, Seung-Gul, and Seo-Eun Cho. 2020. "Neuroimaging Biomarkers for Predicting Treatment Response and Recurrence of Major Depressive Disorder." *International Journal of Molecular Sciences* 21 (6). https://doi.org/10.3390/ijms21062148.

Kass, Robert E., Brian S. Caffo, Marie Davidian, Xiao-Li Meng, Bin Yu, and Nancy Reid. 2016. "Ten Simple Rules for Effective Statistical Practice." *PLOS Computational Biology* 12 (6): e1004961. https://doi.org/10.1371/journal.pcbi.1004961.

Kennedy, Sidney H., Raymond W. Lam, Susan Rotzinger, Roumen V. Milev, Pierre Blier, Jonathan Downar, Kenneth R. Evans, Faranak Farzan, Jane A. Foster, and Benicio N. Frey. 2019. "Symptomatic and Functional Outcomes and Early Prediction of Response to Escitalopram Monotherapy and Sequential Adjunctive Aripiprazole Therapy in Patients With Major Depressive Disorder: A CAN-BIND-1 Report." *The Journal of Clinical Psychiatry* 80 (2): 0–0. https://doi.org/10.4088/JCP.18m12202.

Koolschijn, P. Cédric M. P., Neeltje E. M. van Haren, Hilleke E. Hulshoff Pol, and René

      S. Kahn. 2008. "Hypothalamus Volume in Twin Pairs Discordant for

      Schizophrenia." *European Neuropsychopharmacology: The Journal of the*

      European College of Neuropsychopharmacology 18 (4): 312–15.

      https://doi.org/10.1016/j.euroneuro.2007.12.004.

Lam, Raymond W., Roumen Milev, Susan Rotzinger, Ana C. Andreazza, Pierre Blier,

      Colleen Brenner, Zafiris J. Daskalakis, et al. 2016. "Discovering Biomarkers for

      Antidepressant Response: Protocol from the Canadian Biomarker Integration

      Network in Depression (CAN-BIND) and Clinical Characteristics of the First

      Patient Cohort." *Bmc Psychiatry* 16 (April): 105. https://doi.org/10.1186/s12888-

      016-0785-x.

Liu, Dawei, Hans J. Johnson, Jeffrey D. Long, Vincent A. Magnotta, and Jane S. Paulsen.

      2014. "The Power-Proportion Method for Intracranial Volume Correction in

      Volumetric Imaging Analysis." *Frontiers in Neuroscience* 8.

      https://doi.org/10.3389/fnins.2014.00356.

MacQueen, Glenda M., Stefanie Hassel, Stephen R. Arnott, Addington Jean, Christopher

      R. Bowie, Signe L. Bray, Andrew D. Davis, et al. 2019. "The Canadian

      Biomarker Integration Network in Depression (CAN-BIND): Magnetic

      Resonance Imaging Protocols." *Journal of Psychiatry & Neuroscience: JPN* 44

      (3): 1–14. https://doi.org/10.1503/jpn.180036.

Malberg, Jessica E., René Hen, and Torsten M. Madsen. 2021. "Adult Neurogenesis and

      Antidepressant Treatment: The Surprise Finding by Ron Duman and the Field 20

Years Later." *Biological Psychiatry*, January.

https://doi.org/10.1016/j.biopsych.2021.01.010.

Markakis, Eleni A., Theo D. Palmer, Lynne Randolph-Moore, Pasko Rakic, and Fred H.

Gage. 2004. "Novel Neuronal Phenotypes from Neural Progenitor Cells."

Journal of Neuroscience 24 (12): 2886–97.

https://doi.org/10.1523/JNEUROSCI.4161-03.2004.

Mehta, Divya, Elise S. Pelzer, Dagmar Bruenig, Bruce Lawford, Sarah McLeay, Charles

P. Morris, John N. Gibson, et al. 2019. "DNA Methylation from Germline Cells

in Veterans with PTSD." *Journal of Psychiatric Research* 116 (September): 42–

50. https://doi.org/10.1016/j.jpsychires.2019.06.001.

Milenkovic, Vladimir M., Evan H. Stanton, Caroline Nothdurfter, Rainer Rupprecht, and

Christian H. Wetzel. 2019. "The Role of Chemokines in the Pathophysiology of

Major Depressive Disorder." *International Journal of Molecular Sciences* 20 (9).

https://doi.org/10.3390/ijms20092283.

Miller, Olivia, Jane Shakespeare-Finch, Dagmar Bruenig, and Divya Mehta. 2020. "DNA

Methylation of NR3C1 and FKBP5 Is Associated with Posttraumatic Stress

Disorder, Posttraumatic Growth, and Resilience." *Psychological Trauma:

Theory, Research, Practice, and Policy* 12 (7): 750–55.

http://dx.doi.org.libaccess.lib.mcmaster.ca/10.1037/tra0000574.

Mora, Cristina, Valentina Zonca, Marco A. Riva, and Annamaria Cattaneo. 2018. "Blood

Biomarkers and Treatment Response in Major Depression." *Expert Review of*

Molecular Diagnostics 18 (6): 513–29.

https://doi.org/10.1080/14737159.2018.1470927.

Mulder, Rosa H., Jolien Rijlaarsdam, Maartje P. C. M. Luijk, Frank C. Verhulst, Janine F.

Felix, Henning Tiemeier, Marian J. Bakermans-Kranenburg, and Marinus H. van

IJzendoorn. 2017. "Methylation Matters: FKBP506 Binding Protein 51 (FKBP5)

Methylation Moderates the Associations of FKBP5 Genotype and Resistant

Attachment with Stress Regulation." *Development and Psychopathology* 29 (2):

491–503. https://doi.org/10.1017/S095457941700013X.

Nogovitsyn, Nikita, Meghan Muller, Roberto Souza, Stefanie Hassel, Stephen R. Arnott,

Andrew D. Davis, Geoffrey B. Hall, et al. 2019. "Hippocampal Tail Volume as a

Predictive Biomarker of Antidepressant Treatment Outcomes in Patients with

Major Depressive Disorder: A CAN-BIND Report." *Neuropsychopharmacology*,

October, 1–9. https://doi.org/10.1038/s41386-019-0542-1.

Pace, Thaddeus W. W., Fang Hu, and Andrew H. Miller. 2007. "Cytokine-Effects on

Glucocorticoid Receptor Function: Relevance to Glucocorticoid Resistance and

the Pathophysiology and Treatment of Major Depression." *Brain, Behavior, and

Immunity* 21 (1): 9–19. https://doi.org/10.1016/j.bbi.2006.08.009.

Perna, Giampaolo, Alessandra Alciati, Silvia Daccò, Massimiliano Grassi, and Daniela

Caldirola. 2020. "Personalized Psychiatry and Depression: The Role of

Sociodemographic and Clinical Variables." *Psychiatry Investigation* 17 (3):

193–206. https://doi.org/10.30773/pi.2019.0289.

Piechaczek, Charlotte Elisabeth, Ellen Greimel, Lisa Feldmann, Verena Pehl, Antje-Kathrin Allgaier, Michael Frey, Franz Joseph Freisleder, et al. 2019. "Interactions between FKBP5 Variation and Environmental Stressors in Adolescent Major Depression." *Psychoneuroendocrinology* 106: 28–37. https://doi.org/10.1016/j.psyneuen.2019.03.025.

Rao, Shuquan, Yao Yao, Joanne Ryan, Tao Li, Duan Wang, Chuan Zheng, Yong Xu, and Qi Xu. 2016. "Common Variants in FKBP5 Gene and Major Depressive Disorder (MDD) Susceptibility: A Comprehensive Meta-Analysis." *Scientific Reports* 6: 32687. https://doi.org/10.1038/srep32687.

Roy, Bhaskar, Richard C. Shelton, and Yogesh Dwivedi. 2017. "DNA Methylation and Expression of Stress Related Genes in PBMC of MDD Patients with and without Serious Suicidal Ideation." *Journal of Psychiatric Research* 89: 115–24. https://doi.org/10.1016/j.jpsychires.2017.02.005.

Santos, Marcelo Antônio Oliveira, Lucas Soares Bezerra, Ana Rita Marinho Ribeiro Carvalho, and Alessandra Mertens Brainer-Lima. 2018. "Global Hippocampal Atrophy in Major Depressive Disorder: A Meta-Analysis of Magnetic Resonance Imaging Studies." *Trends in Psychiatry and Psychotherapy* 40 (4): 369–78. https://doi.org/10.1590/2237-6089-2017-0130.

Schindler, S., L. Schmidt, M. Stroske, M. Storch, A. Anwander, R. Trampel, M. Strauß, U. Hegerl, S. Geyer, and P. Schönknecht. 2019. "Hypothalamus Enlargement in Mood Disorders." *Acta Psychiatrica Scandinavica* 139 (1): 56–67. https://doi.org/10.1111/acps.12958.

Schindler, Stephanie, Stefan Geyer, Maria Strauß, Alfred Anwander, Ulrich Hegerl, Robert Turner, and Peter Schönknecht. 2012. "Structural Studies of the Hypothalamus and Its Nuclei in Mood Disorders." *Psychiatry Research: Neuroimaging* 201 (1): 1–9. https://doi.org/10.1016/j.pscychresns.2011.06.005.

Schindler, Stephanie, Peter Schönknecht, Laura Schmidt, Alfred Anwander, Maria Strauß, Robert Trampel, Pierre-Louis Bazin, et al. 2013. "Development and Evaluation of an Algorithm for the Computer-Assisted Segmentation of the Human Hypothalamus on 7-Tesla Magnetic Resonance Images." *PLOS ONE* 8 (7): e66394. https://doi.org/10.1371/journal.pone.0066394.

Stetler, Cinnamon, and Gregory E. Miller. 2011. "Depression and Hypothalamic-Pituitary-Adrenal Activation: A Quantitative Summary of Four Decades of Research." *Psychosomatic Medicine* 73 (2): 114–26. https://doi.org/10.1097/PSY.0b013e31820ad12b.

Suh, Jee Su, Luciano Minuzzi, Pradeep Reddy Raamana, Andrew Davis, Geoffrey B. Hall, Jacqueline Harris, Stefanie Hassel, et al. 2020. "An Investigation of Cortical Thickness and Antidepressant Response in Major Depressive Disorder: A CAN-BIND Study Report." *NeuroImage : Clinical* 25 (January). https://doi.org/10.1016/j.nicl.2020.102178.

Terlevic, Robert, Miriam Isola, Maria Ragogna, Martina Meduri, Francesca Canalaz, Laura Perini, Gianluca Rambaldelli, et al. 2013. "Decreased Hypothalamus Volumes in Generalized Anxiety Disorder but Not in Panic Disorder." *Journal of Affective Disorders* 146 (3): 390–94. https://doi.org/10.1016/j.jad.2012.09.024.

Tozzi, Leonardo, Chloe Farrell, Linda Booij, Kelly Doolin, Zsofia Nemoda, Moshe Szyf, Florence B Pomares, Julian Chiarella, Veronica O'Keane, and Thomas Frodl. 2018. "Epigenetic Changes of FKBP5 as a Link Connecting Genetic and Environmental Risk Factors with Structural and Functional Brain Changes in Major Depression." *Neuropsychopharmacology* 43 (5): 1138–45. https://doi.org/10.1038/npp.2017.290.

Vreeburg, Sophie A., Witte J. G. Hoogendijk, Johannes van Pelt, Roel H. DeRijk, Jolanda C. M. Verhagen, Richard van Dyck, Johannes H. Smit, Frans G. Zitman, and Brenda W. J. H. Penninx. 2009. "Major Depressive Disorder and Hypothalamic-Pituitary-Adrenal Axis Activity: Results From a Large Cohort Study." *Archives of General Psychiatry* 66 (6): 617–26. https://doi.org/10.1001/archgenpsychiatry.2009.50.

Wolff, Julia, Stephanie Schindler, Christian Lucas, Anne-Sophie Binninger, Luise Weinrich, Jan Schreiber, Ulrich Hegerl, et al. 2018. "A Semi-Automated Algorithm for Hypothalamus Volumetry in 3 Tesla Magnetic Resonance Images." *Psychiatry Research: Neuroimaging* 277 (July): 45–51. https://doi.org/10.1016/j.pscychresns.2018.04.007.

Xie, Yinping, Ling Xiao, Lijuan Chen, Yage Zheng, Caixia Zhang, and Gaohua Wang. 2021. "Integrated Analysis of Methylomic and Transcriptomic Data to Identify Potential Diagnostic Biomarkers for Major Depressive Disorder." *Genes* 12 (2). https://doi.org/10.3390/genes12020178.

Zou, Hui, and Trevor Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2): 301–20. https://doi.org/10.1111/j.1467-9868.2005.00503.x.

Zuo, Zhiwei, Shuhua Ran, Yao Wang, Chang Li, Qi Han, Qianying Tang, Wei Qu, and Haitao Li. 2019. "Asymmetry in Cortical Thickness and Subcortical Volume in Treatment-Naïve Major Depressive Disorder." *NeuroImage. Clinical* 21: 101614. https://doi.org/10.1016/j.nicl.2018.101614.

5.9 Figures and Tables
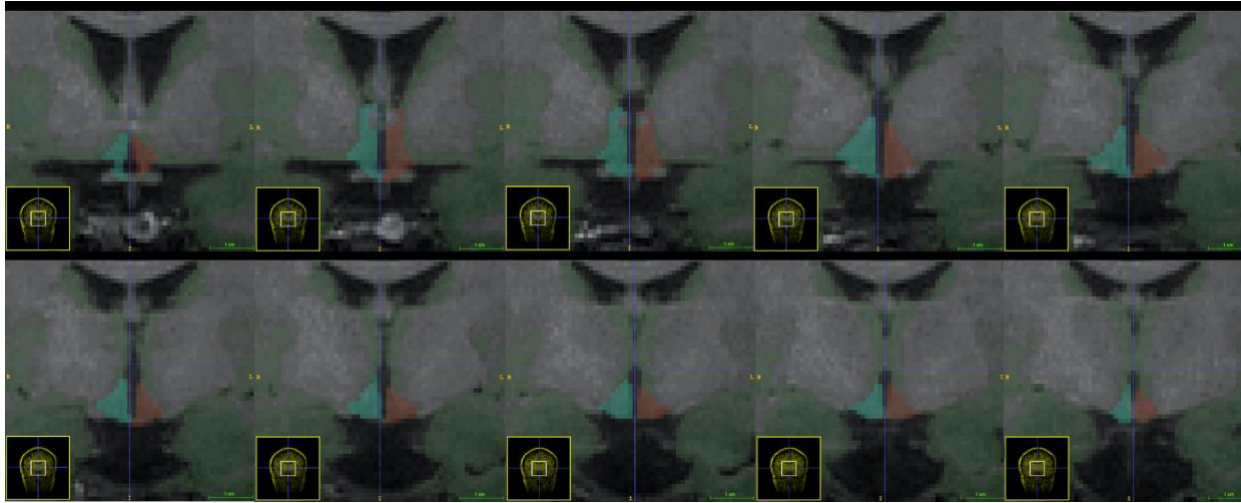
Figure 1**. Representative view of a manually segmented hypothalamus (left = red, right = blue).** Panels viewed from left to right in each row represent 8 contiguous slices, mammillary bodies not included for space. The image belongs to a HC participant.
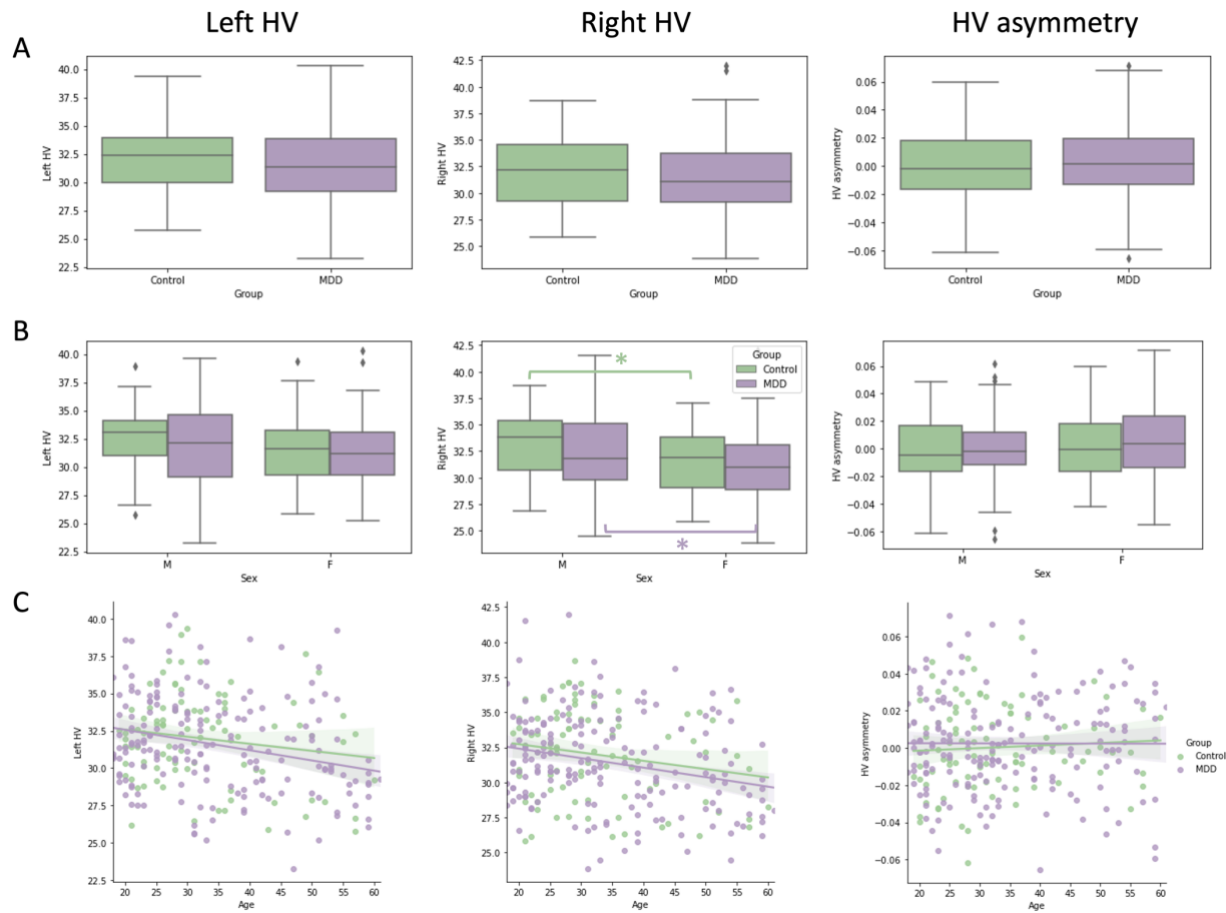
Figure 2. A) Boxplots illustrating the difference in volume measures between HC and MDD. B) Boxplots illustrating the difference in volume measures between males and females, grouped by diagnostic status in each hemisphere. Asterisks (*) indicate that right HV exhibited a significant sex difference (p=0.05) in volume in both MDD (purple) and HC (green). Gray diamonds indicate datapoints that lie outside 1.5 interquartile range above and below the median. C) Scatterplots illustrating

correlations between left and right HV and age, grouped by diagnostic status. Correlations were significant only in left HV and right HV in the MDD group.
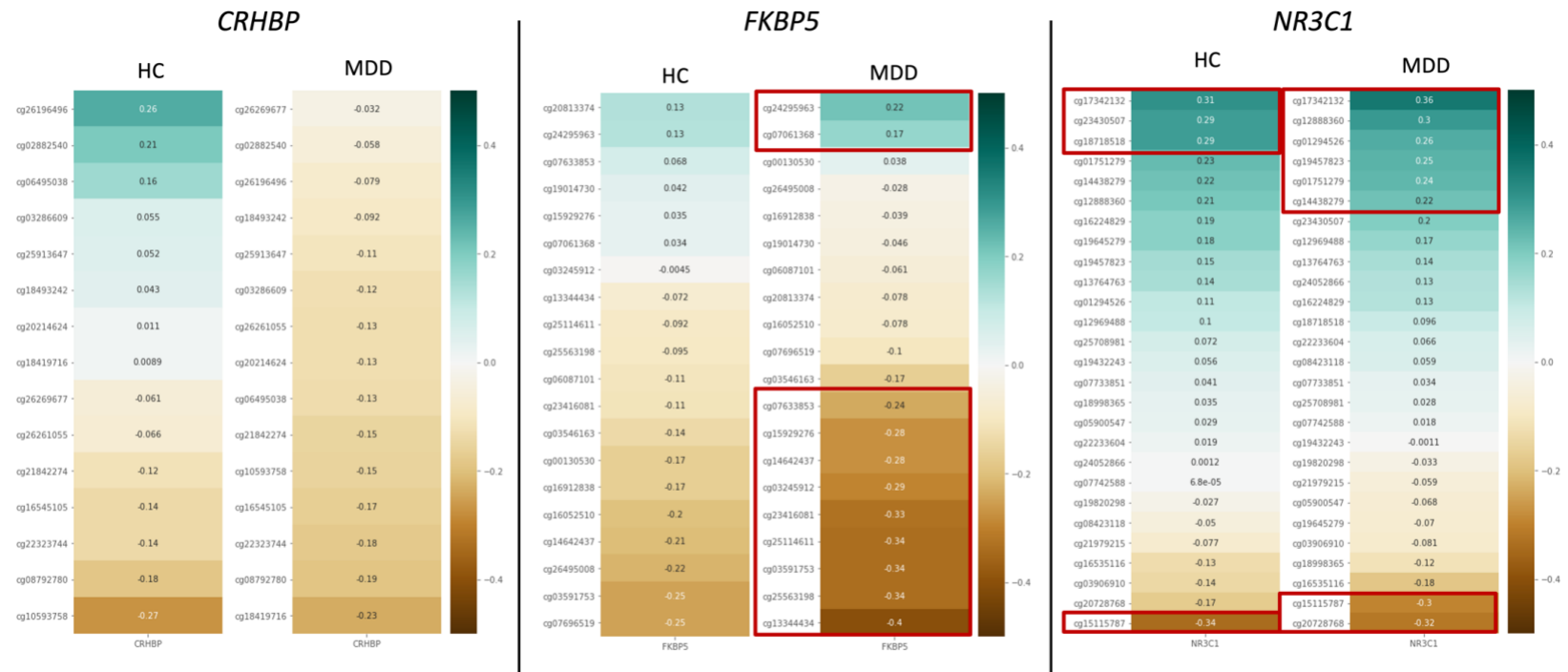
Figure 3. A) Heatmaps illustrating the correlation of DNA methylation at each CpG site with gene expression for each of the candidate genes. The CpG sites whose correlations with gene expression were significant are enclosed in the red boxes. The MDD group exhibited a significantly higher ratio of correlated CpG sites compared to HC in *FKBP5* and *NR3C1*.
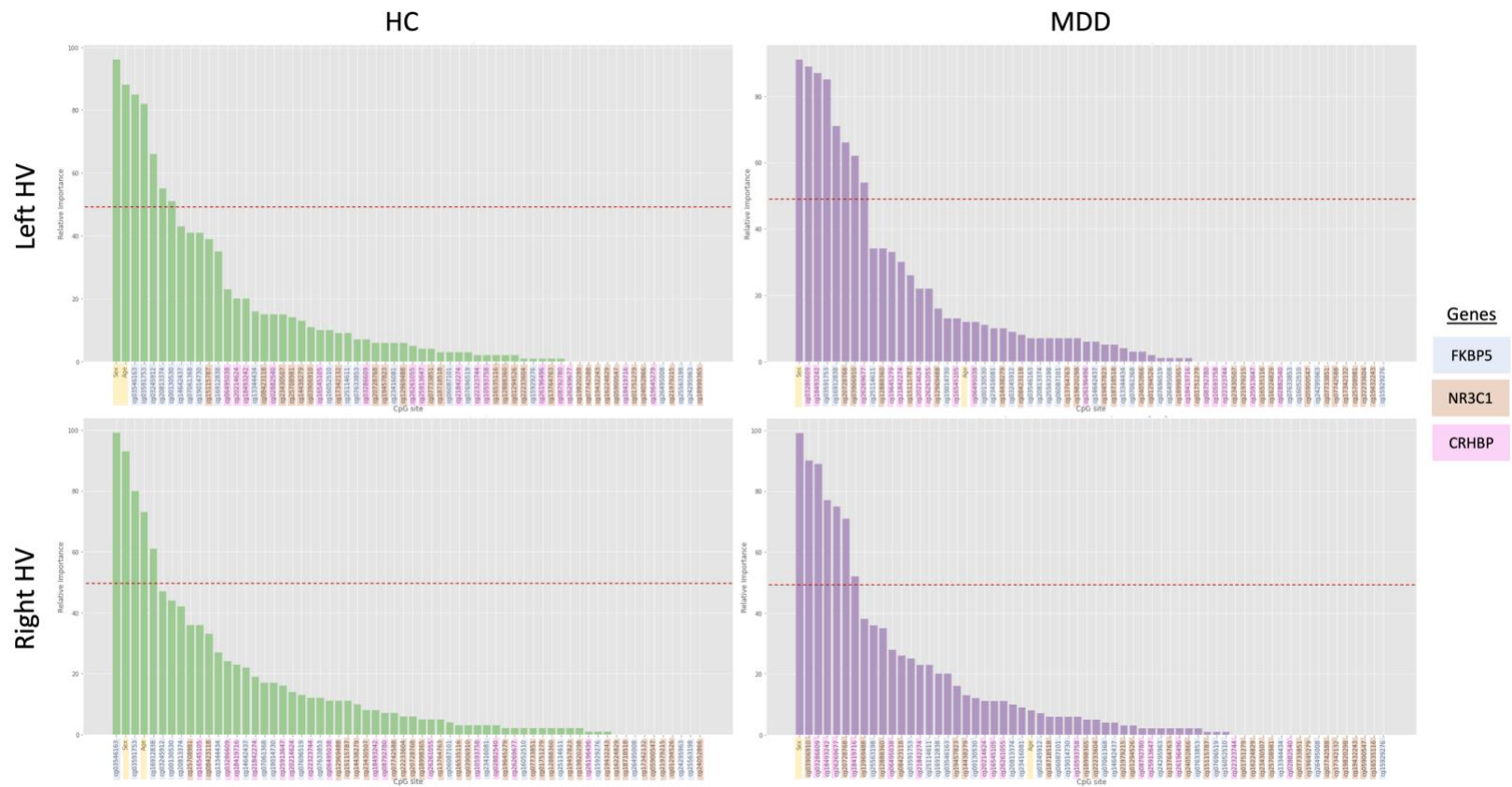
Figure 4. Plots illustrating the RVI scores of the 64 explanatory variables for each elastic net model, encompassing all CpG sites as well as sex and age as reference variables. Variables are colour coded to indicate the gene on which they are located as illustrated by the legend. The threshold for a CpG site being considered a highly ranked variable is denoted with a dashed red line (RVI=50). See Table 2 for the model mean error/fit parameters and the top-ranked variables for each model.

Table 1. Demographic and clinical information for the MDD and HC samples.

	MDD Patients			Healthy Controls			MDD vs HC	
	N	Mean/ frequency	SD	N	Mean/ frequency	SD	Test statistic	p-value
Age $^{(ns)}$	181	34.55	12.30	90	32.66	10.55	t=1.25	0.21
Sex % F $^{(ns)}$	181	65.75	--	90	64.44	--	χ^2=0.006	0.94
Years of Education $^{(***)}$	181	16.86	2.14	90	18.54	2.13	t=-6.11	3.5e-9
Baseline MADRS $^{(***)}$	181	29.78	5.52	90	0.84	1.81	t=48.4	9.1e-135
% Family Hx psych. illness $^{(***)}$	180	78.89	--	86	20.93	--	χ^2=88.3	5.6e-21
Age of Illness Onset	175	20.30	10.05	--	--	--	--	--
Number of previous MDEs	129	3.96	2.62	--	--	--	--	--
% Δ MADRS @ 8 weeks	157	45.60	32.34	--	--	--	--	--
% Δ MADRS @ 16 weeks	145	64.47	28.10	--	--	--	--	--
% Responders @ 8 weeks	157	46.50	--	--	--	--	--	--
% Responders @ 16 weeks	145	72.41	--	--	--	--	--	--
CECA score $^{(***)}$	159	6.12	2.89	82	4.20	1.98	χ^2=21.5	3.6e-6
Ethnicity	181	--	--	90	--	--	--	--
Caucasian	121	66.30	--	59	65.56	--	χ^2=0.015	0.90
Asian	26	25.01	--	19	21.11	--	χ^2=1.22	0.27
Hispanic	10	5.52	--	2	2.22	--	χ^2=1.54	0.21
Black	5	2.76	--	2	2.22	--	χ^2=0.069	0.79
Other	18	9.94	--	5	5.56	--	χ^2=1.48	0.22
Prefer no answer	1	0.56	--	2	2.22	--	χ^2=1.50	0.22

Superscripts indicate the significance of the test statistic comparing patient and healthy control samples. 'ns' – p > 0.05, no significant differences between samples. '***' – p < 0.005. The N indicates the number of participants for which the corresponding information is available.

Table 2. Mean elastic net model error parameters (out of 100 runs) and top-ranking CpG sites (RVI>50) for each model, listed in order of RVI/

Group	HV measure	MSE (std)	MAE (std)	R^2 (std)	Variables	Gene	RVI	Coefficient sign
MDD	RHV	9.94 (1.74)	2.54 (0.24)	0.11 (0.07)	Sex	--	99	–
					cg03906910	*NR3C1*	90	+
					cg03286609	*CRHBP*	89	+
					cg18493242	*CRHBP*	77	+
					cg26269677	*CRHBP*	75	+
					cg20728768**	*NR3C1*	71	+
					cg18419716	*CRHBP*	52	+
	LHV	9.34 (1.83)	2.41 (0.26)	0.10 (0.07)	Sex	--	91	–
					cg03286609	*CRHBP*	89	+
					cg18493242	*CRHBP*	87	+
					cg03591753**	*FKBP5*	85	+
					cg16912838	*FKBP5*	71	+
					cg20728768**	*NR3C1*	66	+
					cg03906910	*NR3C1*	62	+
					cg26269677	*CRHBP*	54	+
HC	RHV	11.13 (2.24)	2.82 (0.32)	0.039 (0.04)	cg03546163	*FKBP5*	99	+
					Sex	--	93	–
					cg03591753	*FKBP5*	80	+
					Age	--	73	–
					cg16912838	*FKBP5*	61	+
	LHV	9.93 (2.36)	2.54 (0.35)	0.021 (0.026)	Sex	--	96	–
					Age	--	88	–
					cg03546163	*FKBP5*	85	+
					cg03591753	*FKBP5*	82	+
					cg03245912	*FKBP5*	66	+
					cg20813374	*FKBP5*	55	–
					cg00130530	*FKBP5*	51	–

Double asterisks (**) indicate CpG sites that were significantly related to respective gene expression levels.

5.10 Supplementary Material

5.10.1 Participants

Participants were recruited from six clinical sites in Canada: the Djavad Mowafaghian Centre for Brain Health (Vancouver, BC), Hotchkiss Brain Institute (Calgary, AB), University Health Network (Toronto, ON), Centre for Addiction and Mental Health (Toronto, ON), St. Joseph's Healthcare Hamilton (Hamilton, ON) and Providence Care Hospital (Kingston, ON). Recruitment of MDD and HC participants involved referrals from outpatient clinics and advertisements in the community. Subjects included in the MDD group were outpatients aged 18-60 who met DSM-IV-TR criteria for a major depressive episode (MDE) in MDD. A confirmation of MDD on the Mini International Neuropsychiatric Interview (MINI) and a score greater than 24 on the Montgomery-Åsberg Depression Scale (MADRS) was used by all clinicians to confirm the MDE. Healthy control subjects were age- and sex-matched and had no history of psychiatric illness. All subjects were fluent in English and free of psychotropic medications for at least 5 half-lives before the baseline visit. Exclusion criteria for MDD subjects were: co-morbid diagnosis of any psychiatric illness considered the primary diagnosis, including bipolar disorder Type I and II or personality disorder that would interfere with participation; high suicidal risk; substance dependence or abuse in past 6 months; neurological or major medical condition; pregnancy or breastfeeding; psychosis in current episode; high risk for hypomanic switch; failure of four or more previous pharmacotherapeutic interventions; previous intolerance to escitalopram or aripiprazole;

having started psychological treatment in past three months with intent to continue and

any contraindications to MRI.

Supplemental Figure 1. Boxplots displaying the distribution of left hypothalamus volume, categorized by site and stratified by diagnostic group, before and after applying the ComBat harmonization procedure. CAM – Centre for Addiction and Mental Health; MCU – McMaster University; QNS – Queen's University; TGH – Toronto General Hospital; UBC – University of British Columbia; UCA – University of Calgary.
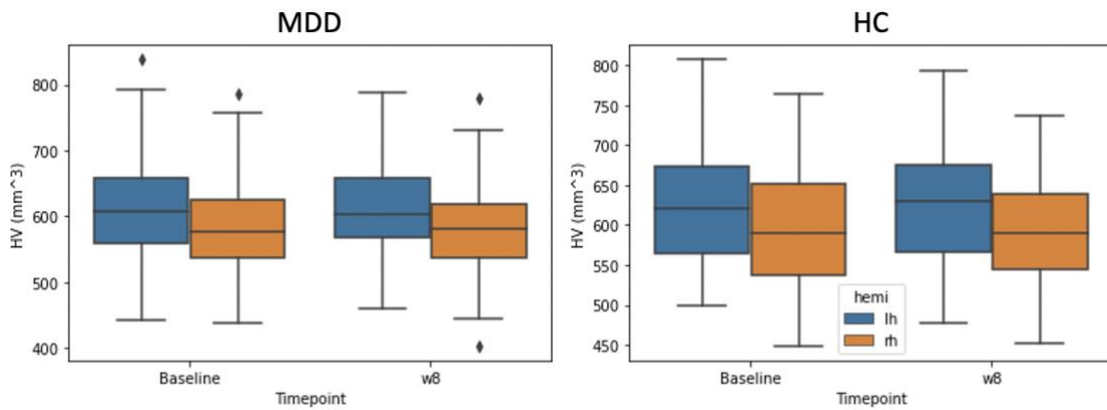
Supplemental Figure 2. Boxplots illustrating change in HV (blue: left HV; orange: right HV) over 8 weeks of escitalopram treatment in the MDD group, compared with the non-treated HC group.

		MDD					HC		Total
Site	*N*	Mean Age	% F	Baseline MADRS	# 16-wk Responders	*N*	Mean Age	%F	Total N
CAM	8	31.25	87.50	26.63	2/5	7	31.00	28.57	15
MCU	27	34.93	70.37	29.56	20/24	16	33.69	68.75	43
QNS	11	40.55	45.45	30.09	6/8	11	29.91	81.82	22
TGH	48	35.50	62.50	32.02	24/37	20	33.95	70.00	68
UBC	59	34.64	67.80	28.03	35/48	9	34.22	77.78	68
UCA	28	30.96	64.29	30.61	18/23	27	32.11	55.56	55

Supplemental Table 1. Demographic and clinical information grouped by site.
MADRS – Montogomery-Asberg Depression Rating Scale

HV	Inter-rater agreement		Average intra-rater agreement	
	Numeric (ICC)	Spatial (Dice)	Numeric (ICC)	Spatial (Dice)
Left	0.85	0.84	0.78	0.91
Right	0.83	0.82	0.79	0.91
Left + Right	0.80	0.83	0.79	0.91

Supplemental Table 2. Numeric and spatial agreement between and within independent raters for manual segmentation of the hypothalamus. Inter-rater agreement was obtained using all 497 segmentations and intra-rater agreement using 50 randomly chosen images.

HV	Mean difference (mm^3)	Test statistic	95% Confidence Interval	Cohen's D	p-value	Power	Ideal sample size
Left	0.41	1.02	[-0.39, 1.2]	0.13	0.314	0.17	930
Right	0.55	1.31	[-0.28, 1.38]	0.168	0.192	0.25	557
Asymmetry	-0.002	-0.69	[-0.01, 0.00]	0.086	0.492	0.10	2123

Supplemental Table 3. Mean differences and associated statistics of t-tests in HV measures between MDD and HC. Results from power analyses are included in the last two columns.

	MDD		HC		MDD vs HC
Gene	CpG site	Location	CpG site	Location	Ratio comparison
CRHBP	--	--	--	--	ns
FKBP5	cg07633853	Body	--	--	$\chi^2 = 77.25$ p<0.0001
	cg13344434	Body			
	cg03591753	5'UTR			
	cg25563198	TSS1500			
	cg15929276	5'UTR			
	cg03245912	TSS1500			
	cg25114611	TSS1500			
	cg07633853	Body			
	cg24295963	5'UTR			
	cg14642437	5'UTR			
	cg23416081	5'UTR			
NR3C1	cg17342132	Body	cg15115787	Body	$\chi^2 = 7.29$ p = 0.007
	cg20728768	Body	cg17342132	Body	
	cg15115787	Body	cg18718518	TSS1500	
	cg12888360	Body	cg23430507	5'UTR	
	cg01751279	5'UTR			
	cg01294526	TSS1500			
	cg14438279	5'UTR			
	cg19457823	Body			

Supplemental Table 4. CpG sites that were significantly correlated with expression levels of their respective genes (within MDD and HC separately) and descriptions of their locations. The last column displays the test statistic and associated p-value of the difference in numbers of significant CpG sites between MDD and HC for each gene.

Chapter 6: Discussion

6.1 Summary of findings

This thesis examined structural neuroimaging correlates of MDD neurobiology and antidepressant response in adults. Anatomical measures, such as cortical thickness and HV, are appealing in their relative ease of measurement and characterization in comparison to functional approaches to neuroimaging, motivating researchers to investigate whether structural features can be of use in predictive models of antidepressant response. Although there have been some promising findings, the majority of these studies suffer from lack of statistical power, insufficient or inappropriate control of confounding variables and/or lack of replication in independent samples, among other limitations. The studies herein suggest that although there are average differences in cortical thickness between those diagnosed with MDD and HC participants, structural neuroimaging may not be an appropriate method to predict response to antidepressants on a short-term period (weeks to months).

Chapter 2 thoroughly characterized the extant literature of cortical thickness in MDD, mostly comprised of cross-sectional comparisons with HC and to a lesser extent, longitudinal studies of treatment-induced changes. The quantitative meta-analysis identified regions of cortical thinning and/or thickening in distinguishing MDD participants from HC using seed-based d mapping. In the whole MDD sample, we observed cortical thinning in bilateral orbitofrontal cortex, left pars opercularis and left calcarine fissure/lingual gyrus. In the subgroup analysis of first episode MDD, we observed thicker supramarginal cortex and thinner left fusiform gyrus. These findings

were further validated by a sensitivity analysis and quantification of publication bias. We critically appraised the literature to highlight the main limitations and make recommendations for future studies.

Chapter 3 was partly motivated by the fact that, although widely conjectured, there was insufficient proof that cortical thickening occurs as a result of antidepressant treatment. This study specifically investigated the effects of a serotonin-norepinephrine reuptake inhibitor (SNRI), desvenlafaxine succinate, on cortical thickness over an 8-week period. We also used the opportunity to run a confirmatory analysis of cortical thickness differences between groups and investigate whether baseline cortical thickness was associated with differential clinical response to DVS. Using FreeSurfer and a full arsenal of quality control (QC) protocols including corrections for multiple comparisons, we found that only non-responders to DVS exhibited higher baseline cortical thickness in the left pars orbitalis compared to HC. We attempted to replicate a previously published result of an association between symptom improvement and baseline cortical thickness values for 5 ROIs comprising frontal, temporal, and anterior cingulate regions, which was not successful, although numerous methodological and sample differences may have contributed to the discrepancy.

Chapter 4 comprised a comprehensive investigation of cortical thickness in a large sample of well-characterized MDD participants. Given all that we knew about prior limitations and constraints in this area of research, it was an important opportunity to confirm what we already suspected and discover new features of interest. Again, using FreeSurfer, we implemented the ENIGMA QC protocol and manual edits of volumes to

maximize data quality and optimize the ensuing cortical thickness measurements. In addition, we implemented a vertex-wise correction for site effects using the ComBat algorithm and accounted for the non-linear effect of age. We replicated the finding of thinner frontal cortex in MDD, specifically in the left rostral middle frontal region. There were no differences in baseline cortical thickness between response groups, no longitudinal changes with treatment, and no relationship with symptom severity.

The approach for Chapter 5 was to combine both neuroimaging and molecular variables bounded within the theoretical framework of HPA axis function. It was the first study to address the relationships between hypothalamus volume (HV), DNA methylation of stress axis genes and the influence of childhood maltreatment. The hypothalamus regulates a range of autonomic and somatic functions including sleep, sexual function, and perhaps most relevant to MDD, the physiological stress response. The main focus was to discern the relationship between HV and gene expression/DNA methylation of *CRHBP, FKBP5* and *NR3C1*. Using blinded manual segmentation, we obtained a measure of HV with good inter-rater reliability. Left HV was negatively correlated with current episode duration, although it was not associated with childhood maltreatment, symptom severity or antidepressant response. We found that the MDD group exhibited a greater number of correlations between the expression and DNA methylation profiles of *NR3C1* and *FKBP5*, indicating greater functional relevance of DNA methylation in these genes. A cross-validated elastic net analysis revealed that DNA methylation explained more variance in HV in the MDD group than in HC; moreover, methylation at two functionally relevant sites positively correlated with HV in MDD but not in HC. Taken

together, these results demonstrated that MDD exhibited greater statistical concordance among biologically disparate variables related to stress. It validates the existence of at least one coordinated neuro-physiological process in the disorder, although the direction of the relationship remains to be seen.

6.2 Significance of overall findings

The results in Chapters 2-5 suggest an effect of left lateralization in MDD with respect to diagnostic status for cortical thickness and symptom severity for HV, which is consistent with what has previously been shown with emotion regulation circuitry and white matter microstructure (M. L. Phillips et al., 2015). Notably, all the significant regions identified in the meta-analysis in Chapter 2 were also located on the left hemisphere (with the exception of thinner orbitofrontal cortex, which occurred on both hemispheres). Currently, findings on brain structural asymmetry are equivocal. Greater right-sided asymmetry in the orbitofrontal cortex was found to correlate with anhedonia in MDD (Dotson et al., 2021), which may correspond to smaller left-sided volumes overall in MDD. However, Zuo et al. (2019) demonstrated that their MDD group exhibited higher left-sided asymmetry in the frontal cortex (including the rostral middle frontal region), corresponding either to smaller right-sided or greater left-sided volumes. A meta-analysis of 59 studies comprising 3280 MDD participants showed that GM decreases in the frontal, temporal, parietal, and limbic lobes were predominant in the right side (Huang et al., 2021). Moreover, a large study (2256 MDD and 3504 HC participants) combining data from 31 sources found no group effect of asymmetry, indicating that there

is no consistent MDD-specific asymmetry in cortical structure, at least in terms of thickness (de Kovel et al., 2019). Positive results from smaller studies may be more reflective of characteristics specific to each cohort or individual differences between participants (more on this below).

For cortical thickness, we replicated the finding that MDD tends to exhibit thinner cortex in prefrontal and frontal regions in medication-free subjects (specifically in Chapters 2 and 4). Most recently, it was observed in a sample of 104 MDD participants that gray matter was reduced in the left rectus and bilateral middle frontal gyri (S. Zhao et al., 2021). The middle frontal gyrus was also implicated among other widespread regions in group-level differences between MDD and HC (Y. Wang et al., 2021). However, most other brain-behaviour or brain-experience associations that have been observed were not replicated by the current studies. For instance, the often cited yet preliminary finding of ACC volume being decreased in those with MDD (Belleau et al., 2019; Ibrahim et al., 2022; Mak et al., 2009) or as being predictive of treatment response (Chen et al., 2007; J. L. Phillips et al., 2015; Sämann et al., 2013) did not materialize as a significant result in this thesis. A recent meta-analysis found that the ACC exhibits a quadratic trajectory across the lifespan, which may partly explain discrepant findings that only account for a linear component of age (Frangou et al., 2022). Another reason may be that studies of the ACC are heavily influenced by theory and pre-existing assumptions about regions putatively involved in emotion regulation, meaning it is often selected for targeted ROI analyses. Thus, any effects in this region may be more likely to be excluded from the

results of whole-brain analyses, where the threshold for statistical significance is much higher.

With respect to treatment response in general, higher frontal volume has previously been associated with response, whereas higher temporal volumes have been associated with non-response, across 5 studies included in a recent review (Enneking et al., 2020). However, 3 of these studies had a sample of less than 45 MDD participants (Chen et al., 2007; Li et al., 2010; Liao et al., 2013), 1 studied a nonlinear relationship between cortical volume and response (Korgaonkar et al., 2015) and 1 study included participants suffering from unipolar and bipolar depression in their sample (Sämann et al., 2013). Likewise, the evidence for volumetric changes in response to antidepressant treatment is relatively poor, with negative findings observed across 7 out of 11 reviewed studies (Enneking et al., 2020), in addition to the negative results in Chapters 3-5. More recent studies have also observed a lack of association with baseline neuroimaging measures and treatment response to medications, specifically with respect to structural brain age (Ballester et al., 2021) and using positron emission tomography (PET) to quantify glucose metabolism (Hill et al., 2021). PET findings in this area, although less numerous, also suffer from a lack of replication, similar to other modalities.

To predict treatment response on the order of weeks to months, it may be that modalities with higher temporal resolutions would exhibit more utility, such as fMRI or EEG. A recent study using resting-state fMRI found that connectivity measures between the salience, executive control and somatomotor networks predicted treatment outcome to one of three common antidepressants (Braund et al., 2021). Similarly, task-based fMRI

measuring reward processing, in conjunction with clinical variables and demographics, achieved an R^2 of 48% in predicting response to sertraline, although R^2 was smaller for predicting response to adjunctive bupropion and placebo (Nguyen et al., 2022). However, clinical data on symptom severity remained the strongest predictors of treatment response.

Research on hypothalamic structure and function is still in early stages, and the study in Chapter 5 was the first to assess the relationship between hypothalamus volume and DNA methylation of stress axis genes. We cautiously interpreted the finding of a positive association between DNA methylation and hypothalamus volume to be a protective factor. Further evidence for this interpretation can be found in Humphreys et al. (2019), where lower levels of DNA methylation in stress axis genes were associated with greater odds of MDD onset in a sample of young girls. Recent research also highlights the putative role of the hypothalamus in mood and anxiety (Modi et al., 2019; D. Wang et al., 2019). On the other hand, a recent review suggested that rather than the hypothalamus, volumes in the hippocampus, amygdala and frontal cortex are more likely to mediate the link between MDD and childhood maltreatment/adverse events (Silva et al., 2021).

Another common hypothesis in the literature is that symptom severity leaves its mark on, or is reflected by, the macrostructural level via cortical thickness or volumetric measures. However, our review in Chapter 2 regarding clinical correlations of cortical thickness in MDD revealed major inconsistencies among findings. In the primary studies in Chapters 3 - 5, there were no associations to be observed between cortical

thickness/HV and symptom severity (or improvement). Either there really are no associations between brain structure and severity of MDD to be observed, or we are not detecting structural changes or symptom severity with enough sensitivity and/or specificity. Evidence suggests some combination of the two possibilities. The ENIGMA meta-analyses for cortical and subcortical features of depression found no associations between symptom severity and subcortical volumes or cortical thickness in adults (Schmaal et al., 2016, 2017). However, attempts to model higher-level associations between symptoms and brain features have been more successful. In one study, after identifying dysphoric and anxio-somatic symptom clusters, the authors showed that models trained on novel individualized functional connectivity-symptom severity maps reliably predicted dysphoria in a validation cohort (Y. Zhao et al., 2022). Notably, this finding was not replicated using a conventional group average approach in the same sample. Generalizability of findings in this arena seems to be a problem elsewhere. For instance, Drysdale et al., (2017) demonstrated a relationship between brain connectivity and clinical symptoms using multivariate methods but has not been replicated since its publication. In one recent study, the authors did not observe any associations between depression or anxiety with independent brain components extracted from structural and functional modalities and in a relatively large sample of 170 participants (Maglanoc et al., 2020). Rather, the strongest associations they observed with brain components were with sex and age.

Taken all together, these results and subsequent studies from other groups suggest that univariate and group-level findings lack the specificity and sensitivity required for

establishing reliable biomarkers of the subjective experience of depression. A paradigm shift in the way we approach the multimodal and high-dimensional data we collect may be required to yield reliable advances in precision psychiatry.

6.3 Strengths

The CAN-BIND analyses (Chapters 4 and 5) benefited from a relatively large sample size for primary studies, counteracting the likelihood of Type II error in using linear models with multiple terms. Multiple sites of data collection across Canada were leveraged to recruit more participants in tandem, synchronized by standardized protocols.

Imaging analyses throughout this thesis underwent extensive QC measures. For CT analyses, a baseline QC framework was adapted from the ENIGMA consortium and further enhanced with in-depth visual inspections as well as blinded manual edits. For the HV analysis, we drew from the literature on semi-automatic protocols to implement a degree of standardization and aids for manual segmentation, including the use of a gray matter tissue probability map overlay to further distinguish hypothalamic boundaries (Wolff et al., 2018). These strategies were used to mitigate any systematic bias from scanning or automatic segmentation error.

Low-powered neuroimaging analyses can lead to inflated estimates of group differences due to sampling variability (Szucs & Ioannidis, 2020). In this thesis, the possibility of Type 1 error was mitigated by applying stringent multiple testing corrections, either via Bonferroni correction, false discovery rate, cluster-based thresholding, permutations, or a combination of the above. Although the number of

positive results were likely reduced as a result of these protocols, it lends a greater degree of reliability and credibility to the results that did survive.

As a given, we accounted for important covariates such as age and sex in all analyses. In Chapter 5 in particular, we characterized the relationships between HV and these covariates in detail to inform future studies. In Chapter 4, we additionally accounted for quadratic age, as it is now well-known that cortical thickness follows a nonlinear trajectory throughout the lifespan (Frangou et al., 2022; Storsve et al., 2014). Site effects presented a more difficult challenge, but we strove for a balanced approach to site correction in this thesis. For the cortical thickness analysis in Chapter 4, the reasons for site differences in cortical thickness measures were not immediately clear, as they could have been due to scanner differences but could also have been influenced by demographic differences between site samples, such as age distribution or sex ratios. We adapted the ComBat algorithm to apply corrections on a vertex-wise basis to eliminate a significant variability bias that had clearly been skewing findings (i.e., common significant differences between response groups and HC were not reflected in the overall group difference, although power should have been sufficient to detect these effects). This led to a biologically and statistically plausible finding of thinner cortex in the left middle frontal gyrus. However, for the HV analysis in Chapter 5, as only one site exhibited significantly different HV values than the others, the site correction was not well-justified. Since results did not change upon site correction, we prioritized the sparse model to avoid introducing potential bias when it was unnecessary.

6.4 Limitations

The limitations of this thesis work span from ground-level methodological considerations such as sample size to more abstract considerations such as inter-individual variation. The sample size of the CAN-BIND cohort was simultaneously a limitation as well as a strength, as the problem of site effects countered the increase in participants to some extent. In a multisite study, we would ideally use all the participants from one site as a holdout set to validate primary findings; in this thesis, we prioritized maximizing sample size in univariate analyses to settle inconsistent findings, as opposed to constructing generalizable models for personalized prediction.

Even as we maximized sample size, there is still the significant possibility of Type II error occurring in these analyses. The signal-to-noise ratios in complex biological data subsuming multiple sources of variation, including MRI and molecular markers, are notoriously low. Accounting for important covariates may have affected the sources of variation in which we are interested, as it has been shown that the effects of demographic variables such as sex and age may eclipse those of disorder-related processes (Maglanoc et al., 2020; Suh et al., 2021). It may also be that the differences in brain structure between MDD and HC or between response groups are not detectable using current standard MRI field strengths (1.5-3T) or popular automatic protocols.

It is often acknowledged that MDD is a heterogenous disorder clinically, and by extension assumed to be so neurobiologically and physiologically as well (Feczko et al., 2019). However, beyond clinical heterogeneity, there is a strong possibility that normal inter-subject variation is likely to be greater than average group differences, particularly

in the frontal, cingulate and temporal regions that are most often implicated in mood disorders (Dinga et al., 2019; Frangou et al., 2022). This natural variability may occlude disorder-specific features that might be relevant on an individual basis to that individual's experience and outcomes. There may be other non-clinical factors that influence brain measures independent of depression that were not considered, such as impaired social functioning (Zhou et al., 2021), exercise (Bashir et al., 2021; Domingos et al., 2021), excessive internet use (Sadeghi et al., 2021) and/or bilingualism (Anderson et al., 2021), to name a few. Structural changes might occur even before an individual experiences much of life: a recent review found that perinatal depression in the mother was associated with structural alterations in the amygdala and fronto-temporal regions in offspring (Cattarinussi et al., 2021). It was also recently observed in a study investigating post-traumatic stress disorder (PTSD) that different symptom clusters mapped onto different cortical thickness and volumetric features, even as overall differences between PTSD and HC groups were non-significant (Crombie et al., 2021). Going forward, we may need to be more attentive to dimensional aspects of MDD, breaking down the clinical diagnosis status often used in univariate analyses to a dimensional profile (based on severity or presence of specific clinical symptoms or domains as per the Research Domain Criteria [Insel et al., 2010]) for each individual that could be used in multivariate approaches in discerning brain-behaviour relationships. Additionally, identification and replication of robust neural correlates with specific symptoms or dimensions would improve interpretability of future findings.

On the clinical side, the main instrument to measure symptom severity within this thesis was the MADRS, a clinician-administered structured interview. Although it is assumed that being clinician-administered, the MADRS is more reliable than self-report measures, it poses several limitations. First, this scale does not take into account atypical symptoms of depression, such as increased appetite/weight as well as increased sleep. The prevalence of the atypical subtype among those diagnosed with MDD has been estimated to be anywhere between 6% in the UK Biobank cohort (Brailean et al., 2020) to 15.3% in a Chinese cohort (Xin et al., 2019), with a mid-level estimate of 10.23% observed in a large cross-sectional survey (Blanco et al., 2012). These studies additionally found that the atypical subtype is associated with female gender, earlier age of onset, greater symptom severity/episode frequency, obesity, psychiatric comorbidities and increased suicidal ideation/attempts, among others (Blanco et al., 2012; Brailean et al., 2020; Xin et al., 2019). As a result, the MADRS may have underestimated symptom severity in those who experienced atypical symptoms of depression. In the ENIGMA study of cortical features in MDD, associations only with the BDI self-report score were detected with respect to surface area of frontal and parietal regions (Schmaal et al., 2017), but not with the HAM-D (clinician-administered), illustrating a possible explanatory gap between self-report and assessor-administered measures. Since measures of antidepressant response were based on the MADRS as well, this may partly explain the lack of association with a variety of brain structural or molecular variables. Future studies might employ both self-report and clinician-observed measures of symptom severity, either by including both in an analysis or by creating a composite measure.

6.5. Future directions

6.5.1 Deciphering the neurobiology of MDD

As mentioned above, there is likely no 1-to-1 correspondence between low-level imaging features (e.g., average cortical thickness values of independent regions) and what we know as 'mood', not to mention all the other somatic, vegetative, and dysphoric symptoms that comprise the MDD diagnosis. Mood on its own is a complex, higher-order process that is likely emergent from the interaction of multiple regions (variable both spatially and temporally) rather than something that can be mapped to immediately detectable features at individual regions (Horien et al., 2020). This highlights the importance and relevance of connection-based measures, which can be applied both structurally and functionally and are also more robust to variations in parcellation (Lord et al., 2016). One example is to employ targeted dimensionality reduction approaches such as independent component analysis to characterize distinct sources of variation that may be distributed spatiotemporally in the brain (Maglanoc et al., 2020). Another approach is to use graph theoretical methods to characterize structural and functional network organization and explicitly model statistical dependence among brain regions (Yun & Kim, 2021). More generally, higher-level, connection-based measures should become the standard in clinical translational efforts going forward, as goals for effective modelling start to outstrip the utility of univariate comparisons and standard parametric mapping.

In addition, socioeconomic factors should be considered with more care and precision. It has been found that not only were race and household income independently associated with likelihood of MDE in 12 months, but gender also interacted with race and income to predict MDE in a sample of over 4000 individuals in the National Survey of American Life (Assari, 2017). A recent review on the biological, psychological, and social determinants of depression illustrated a complex landscape of interacting risk and protective factors at the individual and societal levels (Remes et al., 2021). Considering the multitude of factors involved in depressive onset, progression, and outcomes, it seems almost self-evident that between-individual variation would be greater than average between-group differences.

Recently, Reddan (2021) recommended a structured approach for developing clinically relevant neuroimaging models of chronic pain that properly reflects the biopsychosocial model. As depression shares similarities with chronic pain in many aspects, we might extrapolate these recommendations to guide our efforts. The approach is divided into 3 levels. The first level is the nomothetic (population-based) approach to diagnosis, where one seeks group differences that would theoretically be generalizable to the larger population—here, researchers strive to "control for" as many individual differences as possible, in order to discern the group-level signal. The finding of thinner cortex in prefrontal areas replicated in this thesis would be an example of a nomothetic finding.

The second level of the approach tackles heterogeneity of the disorder, specifically the fact that even a robust group difference may not be reflected in a given

individual. This approach essentially involves training a model on many samples from the same individual. Due to cost and time requirements of the scanning that would be involved, a regularized model can be used that incorporates group-level priors to considerably reduce data acquisition demands (Lindquist et al., 2017). Finally, socioeconomic diversity can be incorporated by thoroughly assessing an individual's SES from different perspectives (including their subjective perception of it) and relating it to brain features. It was suggested to create a thorough SES survey that spans personality, external conditions, and internalization or beliefs of external conditions (Reddan, 2021). At the very least, we may be able to construct a brain component that reflects SES, to be incorporated in higher-level predictive models.

6.5.2 Predicting treatment response in MDD

Given the ambiguity around the neurobiology of MDD as described by MRI, it is perhaps no surprise that the mission of using MRI to predict treatment response in MDD is still in early stages. At this point, we might take a step back to consider the various factors that may affect prediction of treatment response. In addition to the differences among imaging modalities, depression treatments comprise pharmacotherapy as well as psychotherapy and brain stimulation techniques such as transcranial magnetic stimulation (TMS), electroconvulsive therapy (ECT) and transcranial direct current stimulation (tDCS). Furthermore, pharmacotherapeutics are not made equal, with different antidepressants ostensibly targeting different neurotransmitter systems. There are general treatment guidelines for when to use which treatment, mostly depending on previous

treatment resistance and symptom severity (Davidson, 2010). The fact that successful treatment is still largely predicated on trial and error highlights the need to focus on first principles; that is, the underlying neurobiological mechanisms by which depression occurs (as discussed above) and how various treatments exert their effects to alleviate these symptoms. Again, a dimensional approach would be a good starting point and might render even structural imaging measures more useful. As a case in point, one study observed that using different structural measures to predict different latent symptom dimensions in the HAM-D after ECT was more successful than trying to find one set of imaging features to predict the total clinical score (Wade et al., 2021).

A recent meta-analysis aimed to quantify the overall classification performance of machine learning models predicting single-subject response to various treatments in MDD (Cohen et al., 2021). With respect to treatments, ECT exhibited higher performance than medications, although the difference was not statistically significant across all treatments. With respect to modality, MRI generally performed better than EEG, with surprisingly no quantifiable difference in performance between sMRI and fMRI studies. However, the authors suggest that this last result may be affected by publication bias, where research groups may have found and published significant results only with fMRI measures, even if they had also tested sMRI measures that were negative and remain unpublished. Cohen et al. (2021) also observed that sample sizes of the reviewed studies tended to be low and ran the risk of overfitting, concluding overall that it is still too early to include MRI in clinical decision making. The current state of research in this area calls into question the value addition of neuroimaging overall; that is, whether adding imaging

variables on top of clinical data significantly improves treatment response prediction relative to the cost and time involved in collecting that data. Furthermore, beyond the ability to predict whether or not someone will respond to a given treatment (a binary response), differential biomarkers that can help select the most effective treatment for a given individual are necessary to achieve true clinical utility. As is apparent, there are a multitude of variables and factors that significantly complicate the overall endeavour.

It is likely that the problem is not only scientific but infrastructural as well. Conventional interventional trials focus on a very narrow area of the entire feature space that make up a clinical phenotype and are typically biased towards the 'typical' clinical trial participant rather than patients in the real world. In presenting a solution that strives to be truly integrative, Dickson et al. (2020) proposed the master observational trial (MOT), a protocol construct combining the high-quality data obtained via interventional trials with lower quality but highly abundant real-world data. Real world data might comprise medical records/charts, insurance claim information, and data obtained via mobile devices, such as momentary time sampling of various self-report or actigraphy measures. The MOT is essentially characterized by the lack of rigid inclusion/exclusion criteria and increased flexibility that allows the addition of new testing protocols and subsidiary interventional trials to answer specific questions, all subsumed under a centralized data organization scheme. It presents a possible solution to capture information on MDD that is truly population-level, which would enable precise, individual-oriented descriptions and predictions, a significant advance towards true precision psychiatry.

6.6 Conclusion

In this thesis, we showed that thinning in the middle frontal gyrus was a reliable and significant marker of MDD diagnosis but did not observe any significant cortical thickness associations with antidepressant response. We observed a stronger concordance between hypothalamus volume and blood-derived epigenetic profiles of stress axis genes in the MDD group that was not observed in healthy controls. Overall, we found that neuroimaging biomarkers of neurobiology are easier to isolate than biomarkers of future antidepressant response. We recommend that future studies model statistical dependence between structural brain measures and consider the complexity of depression subtypes and socioeconomic backgrounds.

6.7 References

Anderson, J. A. E., Grundy, J. G., Grady, C. L., Craik, F. I. M., & Bialystok, E. (2021). Bilingualism contributes to reserve and working memory efficiency: Evidence from structural and functional neuroimaging. *Neuropsychologia*, *163*, 108071. https://doi.org/10.1016/j.neuropsychologia.2021.108071

Assari, S. (2017). Social Determinants of Depression: The Intersections of Race, Gender, and Socioeconomic Status. *Brain Sciences*, *7*(12), 156. https://doi.org/10.3390/brainsci7120156

Ballester, P. L., Suh, J. S., Nogovitsyn, N., Hassel, S., Strother, S. C., Arnott, S. R., Minuzzi, L., Sassi, R. B., Lam, R. W., Milev, R., Müller, D. J., Taylor, V. H., Kennedy, S. H., & Frey, B. N. (2021). Accelerated brain aging in major depressive disorder and antidepressant treatment response: A CAN-BIND report. *NeuroImage: Clinical*, *32*, 102864. https://doi.org/10.1016/j.nicl.2021.102864

Bashir, S., Al-Sultan, F., Jamea, A. A., Almousa, A., Alzahrani, M. S., Alhargan, F. A., Abualait, T., & Yoo, W. K. (2021). Physical exercise and cortical thickness in healthy controls: A pilot study. *European Review for Medical and Pharmacological Sciences*, *25*(23), 7375–7379. https://doi.org/10.26355/eurrev_202112_27432

Belleau, E. L., Treadway, M. T., & Pizzagalli, D. A. (2019). The Impact of Stress and Major Depressive Disorder on Hippocampal and Medial Prefrontal Cortex Morphology. *Biological Psychiatry*, *85*(6), 443–453. https://doi.org/10.1016/j.biopsych.2018.09.031

Blanco, C., Vesga-López, O., Stewart, J. W., Liu, S.-M., Grant, B. F., & Hasin, D. S. (2012). Prevalence, correlates, comorbidity and treatment-seeking among individuals with a lifetime major depressive episode with and without atypical features: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry*, *73*(2), 224–232. https://doi.org/10.4088/JCP.10m06227

Brailean, A., Curtis, J., Davis, K., Dregan, A., & Hotopf, M. (2020). Characteristics, comorbidities, and correlates of atypical depression: Evidence from the UK Biobank Mental Health Survey. *Psychological Medicine*, *50*(7), 1129–1138. https://doi.org/10.1017/S0033291719001004

Braund, T. A., Breukelaar, I. A., Griffiths, K., Tillman, G., Palmer, D. M., Bryant, R., Phillips, M. L., Harris, A. W. F., & Korgaonkar, M. S. (2021). Intrinsic Functional Connectomes Characterize Neuroticism in Major Depressive Disorder and Predict Antidepressant Treatment Outcomes. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, S2451-9022(21)00204-4. https://doi.org/10.1016/j.bpsc.2021.07.010

Cattarinussi, G., Aarabi, M. H., Sanjari Moghaddam, H., Homayoun, M., Ashrafi, M., Soltanian-Zadeh, H., & Sambataro, F. (2021). Effect of parental depressive symptoms on offspring's brain structure and function: A systematic review of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *131*, 451–465. https://doi.org/10.1016/j.neubiorev.2021.09.046

Chen, C.-H., Ridler, K., Suckling, J., Williams, S., Fu, C. H. Y., Merlo-Pich, E., & Bullmore, E. (2007). Brain imaging correlates of depressive symptom severity and predictors of symptom improvement after antidepressant treatment. *Biological Psychiatry*, *62*(5), 407–414. https://doi.org/10.1016/j.biopsych.2006.09.018

Cohen, S. E., Zantvoord, J. B., Wezenberg, B. N., Bockting, C. L. H., & van Wingen, G. A. (2021). Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: A systematic review and meta-analysis. *Translational Psychiatry*, *11*, 168. https://doi.org/10.1038/s41398-021-01286-x

Crombie, K. M., Ross, M. C., Letkiewicz, A. M., Sartin-Tarm, A., & Cisler, J. M. (2021). Differential relationships of PTSD symptom clusters with cortical thickness and grey matter volumes among women with PTSD. *Scientific Reports*, *11*(1), 1825. https://doi.org/10.1038/s41598-020-80776-2

Davidson, J. R. T. (2010). Major Depressive Disorder Treatment Guidelines in America and Europe. *The Journal of Clinical Psychiatry*, *71*(suppl E1), 27767. https://doi.org/10.4088/JCP.9058se1c.04gry

de Kovel, C. G. F., Aftanas, L., Aleman, A., Alexander-Bloch, A. F., Baune, B. T., Brack, I., Bülow, R., Busatto Filho, G., Carballedo, A., Connolly, C. G., Cullen, K. R., Dannlowski, U., Davey, C. G., Dima, D., Dohm, K., Erwin-Grabner, T., Frodl, T., Fu, C. H. Y., Hall, G. B., … Francks, C. (2019). No Alterations of Brain Structural Asymmetry in Major Depressive Disorder: An ENIGMA Consortium

Analysis. *American Journal of Psychiatry*, *176*(12), 1039–1049.

https://doi.org/10.1176/appi.ajp.2019.18101144

Dickson, D., Johnson, J., Bergan, R., Owens, R., Subbiah, V., & Kurzrock, R. (2020).

The Master Observational Trial: A New Class of Master Protocol to Advance

Precision Medicine. *Cell*, *180*(1), 9–14.

https://doi.org/10.1016/j.cell.2019.12.009

Dinga, R., Schmaal, L., Penninx, B. W. J. H., van Tol, M. J., Veltman, D. J., van Velzen,

L., Mennes, M., van der Wee, N. J. A., & Marquand, A. F. (2019). Evaluating

the evidence for biotypes of depression: Methodological replication and

extension of. *NeuroImage : Clinical*, *22*, 101796.

https://doi.org/10.1016/j.nicl.2019.101796

Domingos, C., Picó-Pérez, M., Magalhães, R., Moreira, M., Sousa, N., Pêgo, J. M., &

Santos, N. C. (2021). Free-Living Physical Activity Measured With a Wearable

Device Is Associated With Larger Hippocampus Volume and Greater Functional

Connectivity in Healthy Older Adults: An Observational, Cross-Sectional Study

in Northern Portugal. *Frontiers in Aging Neuroscience*, *13*, 729060.

https://doi.org/10.3389/fnagi.2021.729060

Dotson, V. M., Taiwo, Z., Minto, L. R., Bogoian, H. R., & Gradone, A. M. (2021).

Orbitofrontal and Cingulate Thickness Asymmetry Associated with Depressive

Symptom Dimensions. *Cognitive, Affective & Behavioral Neuroscience*, *21*(6),

1297–1305. https://doi.org/10.3758/s13415-021-00923-8

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R. N., Zebley, B., Oathes, D. J., Etkin, A., Schatzberg, A. F., Sudheimer, K., Keller, J., Mayberg, H. S., Gunning, F. M., Alexopoulos, G. S., Fox, M. D., Pascual-Leone, A., Voss, H. U., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28–38. https://doi.org/10.1038/nm.4246

Enneking, V., Leehr, E. J., Dannlowski, U., & Redlich, R. (2020). Brain structural effects of treatments for depression and biomarkers of response: A systematic review of neuroimaging studies. *Psychological Medicine*, *50*(2), 187–209. https://doi.org/10.1017/S0033291719003660

Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A. M., Nigg, J. T., & Fair, D. A. (2019). The Heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends in Cognitive Sciences*, *23*(7), 584–601. https://doi.org/10.1016/j.tics.2019.03.009

Frangou, S., Modabbernia, A., Williams, S. C. R., Papachristou, E., Doucet, G. E., Agartz, I., Aghajani, M., Akudjedu, T. N., Albajes-Eizagirre, A., Alnæs, D., Alpert, K. I., Andersson, M., Andreasen, N. C., Andreassen, O. A., Asherson, P., Banaschewski, T., Bargallo, N., Baumeister, S., Baur-Streubel, R., … Dima, D. (2022). Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3–90 years. *Human Brain Mapping*, *43*(1), 431–451. https://doi.org/10.1002/hbm.25364

Hill, K. R., Gardus, J. D., Bartlett, E. A., Perlman, G., Parsey, R. V., & DeLorenzo, C. (2021). Measuring brain glucose metabolism in order to predict response to antidepressant or placebo: A randomized clinical trial. *NeuroImage. Clinical*, *32*, 102858.

Horien, C., Greene, A. S., Constable, R. T., & Scheinost, D. (2020). Regions and Connections: Complementary Approaches to Characterize Brain Organization and Function. *The Neuroscientist*, *26*(2), 117–133. https://doi.org/10.1177/1073858419860115

Huang, K., Kang, Y., Wu, Z., Wang, Y., Cai, S., & Huang, L. (2021). Asymmetrical alterations of grey matter among psychiatric disorders: A systematic analysis by voxel-based activation likelihood estimation. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *110*, 110322. https://doi.org/10.1016/j.pnpbp.2021.110322

Humphreys, K. L., Moore, S. R., Davis, E. G., MacIsaac, J. L., Lin, D. T. S., Kobor, M. S., & Gotlib, I. H. (2019). DNA methylation of HPA-axis genes and the onset of major depressive disorder in adolescent girls: A prospective analysis. *Translational Psychiatry*, *9*, 245. https://doi.org/10.1038/s41398-019-0582-7

Ibrahim, H. M., Kulikova, A., Ly, H., Rush, A. J., & Sherwood Brown, E. (2022). Anterior cingulate cortex in individuals with depressive symptoms: A structural MRI study. *Psychiatry Research: Neuroimaging*, *319*, 111420. https://doi.org/10.1016/j.pscychresns.2021.111420

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry*, *167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Korgaonkar, M. S., Rekshan, W., Gordon, E., Rush, A. J., Williams, L. M., Blasey, C., & Grieve, S. M. (2015). Magnetic Resonance Imaging Measures of Brain Structure to Predict Antidepressant Treatment Outcome in Major Depressive Disorder. *EBioMedicine*, *2*(1), 37–45. https://doi.org/10.1016/j.ebiom.2014.12.002

Li, C.-T., Lin, C.-P., Chou, K.-H., Chen, I.-Y., Hsieh, J.-C., Wu, C.-L., Lin, W.-C., & Su, T.-P. (2010). Structural and cognitive deficits in remitting and non-remitting recurrent depression: A voxel-based morphometric study. *NeuroImage*, *50*(1), 347–356. https://doi.org/10.1016/j.neuroimage.2009.11.021

Li, J., Chen, J., Kong, W., Li, X., & Hu, B. (2022). Abnormal core functional connectivity on the pathology of MDD and antidepressant treatment: A systematic review. *Journal of Affective Disorders*, *296*, 622–634. https://doi.org/10.1016/j.jad.2021.09.074

Liao, Y.-L., Wang, P.-S., Lu, C.-F., Hung, C.-I., Li, C.-T., Lin, C.-P., Hsieh, J.-C., Su, T.-P., & Wu, Y.-T. (2013). Cortical Shape and Curvedness Analysis of Structural Deficits in Remitting and Non-Remitting Depression. *PLOS ONE*, *8*(7), e68625. https://doi.org/10.1371/journal.pone.0068625

Lindquist, M. A., Krishnan, A., López-Solà, M., Jepma, M., Woo, C.-W., Koban, L., Roy, M., Atlas, L. Y., Schmidt, L., Chang, L. J., Reynolds Losin, E. A., Eisenbarth,

H., Ashar, Y. K., Delk, E., & Wager, T. D. (2017). Group-regularized individual prediction: Theory and application to pain. *NeuroImage*, *145*, 274–287. https://doi.org/10.1016/j.neuroimage.2015.10.074

Lord, A., Ehrlich, S., Borchardt, V., Geisler, D., Seidel, M., Huber, S., Murr, J., & Walter, M. (2016). Brain parcellation choice affects disease-related topology differences increasingly from global to local network levels. *Psychiatry Research: Neuroimaging*, *249*, 12–19. https://doi.org/10.1016/j.pscychresns.2016.02.001

Maglanoc, L. A., Kaufmann, T., Jonassen, R., Hilland, E., Beck, D., Landrø, N. I., & Westlye, L. T. (2020). Multimodal fusion of structural and functional brain imaging in depression using linked independent component analysis. *Human Brain Mapping*, *41*(1), 241–255. https://doi.org/10.1002/hbm.24802

Mak, A. K. Y., Wong, M. M. C., Han, S., & Lee, T. M. C. (2009). Gray matter reduction associated with emotion regulation in female outpatients with major depressive disorder: A voxel-based morphometry study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *33*(7), 1184–1190. https://doi.org/10.1016/j.pnpbp.2009.06.025

Modi, S., Thaploo, D., Kumar, P., & Khushu, S. (2019). Individual differences in trait anxiety are associated with gray matter alterations in hypothalamus: Preliminary neuroanatomical evidence. *Psychiatry Research. Neuroimaging*, *283*, 45–54. https://doi.org/10.1016/j.pscychresns.2018.11.008

Phillips, J. L., Batten, L. A., Tremblay, P., Aldosary, F., & Blier, P. (2015). A
    Prospective, Longitudinal Study of the Effect of Remission on Cortical
    Thickness and Hippocampal Volume in Patients with Treatment-Resistant
    Depression. *International Journal of Neuropsychopharmacology*, *18*(8).
    https://doi.org/10.1093/ijnp/pyv037

Phillips, M. L., Chase, H. W., Sheline, Y. I., Etkin, A., Almeida, J. R. C., Deckersbach,
    T., & Trivedi, M. H. (2015). Identifying Predictors, Moderators, and Mediators
    of Antidepressant Response in Major Depressive Disorder: Neuroimaging
    Approaches. *The American Journal of Psychiatry*, *172*(2), 124–138.
    https://doi.org/10.1176/appi.ajp.2014.14010076

Reddan, M. C. (2021). Recommendations for the Development of Socioeconomically-
    Situated and Clinically-Relevant Neuroimaging Models of Pain. *Frontiers in
    Neurology*, *12*, 700833. https://doi.org/10.3389/fneur.2021.700833

Remes, O., Mendes, J. F., & Templeton, P. (2021). Biological, Psychological, and Social
    Determinants of Depression: A Review of Recent Literature. *Brain Sciences*,
    11(12), 1633. https://doi.org/10.3390/brainsci11121633

Sadeghi, S., Takeuchi, H., Shalani, B., Taki, Y., Nouchi, R., Yokoyama, R., Kotozaki, Y.,
    Nakagawa, S., Sekiguchi, A., Iizuka, K., Hanawa, S., Araki, T., Miyauchi, C.
    M., Sakaki, K., Nozawa, T., Ikeda, S., Yokota, S., Magistro, D., Sassa, Y., &
    Kawashima, R. (2021). Brain structures and activity during a working memory
    task associated with internet addiction tendency in young adults: A large sample

study. *PloS One*, *16*(11), e0259259.

https://doi.org/10.1371/journal.pone.0259259

Sämann, P. G., Höhn, D., Chechko, N., Kloiber, S., Lucae, S., Ising, M., Holsboer, F., &

Czisch, M. (2013). Prediction of antidepressant treatment response from gray

matter volume across diagnostic categories. *European*

Neuropsychopharmacology: The Journal of the European College of

Neuropsychopharmacology, *23*(11), 1503–1515.

https://doi.org/10.1016/j.euroneuro.2013.07.004

Schmaal, L., Hibar, D. P., Saemann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N.,

Cheung, J. W., Erp, T. G. M. van, Bos, D., Ikram, M. A., Vernooij, M. W.,

Niessen, W. J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H. J., Janowitz,

D., Buelow, R., Selonke, M., … Depressive, E. M. (2017). Cortical

abnormalities in adults and adolescents with major depression based on brain

scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder

Working Group. *Molecular Psychiatry*, *22*(6), 900–909.

https://doi.org/10.1038/mp.2016.60

Schmaal, L., Veltman, D. J., van Erp, T. G. M., Sämann, P. G., Frodl, T., Jahanshad, N.,

Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W. J., Vernooij, M. W., Ikram,

M. A., Wittfeld, K., Grabe, H. J., Block, A., Hegenscheid, K., Völzke, H.,

Hoehn, D., Czisch, M., … Hibar, D. P. (2016). Subcortical brain alterations in

major depressive disorder: Findings from the ENIGMA Major Depressive

Disorder working group. *Molecular Psychiatry*, *21*(6), 806–812.

https://doi.org/10.1038/mp.2015.69

Silva, R. C., Maffioletti, E., Gennarelli, M., Baune, B. T., & Minelli, A. (2021).

Biological correlates of early life stressful events in major depressive disorder.

Psychoneuroendocrinology, *125*, 105103.

https://doi.org/10.1016/j.psyneuen.2020.105103

Storsve, A. B., Fjell, A. M., Tamnes, C. K., Westlye, L. T., Overbye, K., Aasland, H. W.,

& Walhovd, K. B. (2014). Differential longitudinal changes in cortical thickness,

surface area and volume across the adult life span: Regions of accelerating and

decelerating change. *The Journal of Neuroscience: The Official Journal of the

Society for Neuroscience*, *34*(25), 8488–8498.

https://doi.org/10.1523/JNEUROSCI.0391-14.2014

Suh, J. S., Fiori, L. M., Ali, M., Harkness, K. L., Ramonas, M., Minuzzi, L., Hassel, S.,

Strother, S. C., Zamyadi, M., Arnott, S. R., Farzan, F., Foster, J. A., Lam, R. W.,

MacQueen, G. M., Milev, R., Müller, D. J., Parikh, S. V., Rotzinger, S., Sassi, R.

B., … Frey, B. N. (2021). Hypothalamus volume and DNA methylation of stress

axis genes in major depressive disorder: A CAN-BIND study report.

Psychoneuroendocrinology, *132*, 105348.

https://doi.org/10.1016/j.psyneuen.2021.105348

Szucs, D., & Ioannidis, J. PA. (2020). Sample size evolution in neuroimaging research:

An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–

2018) in high-impact journals. *NeuroImage*, *221*, 117164. https://doi.org/10.1016/j.neuroimage.2020.117164

Wade, B. S. C., Hellemann, G., Espinoza, R. T., Woods, R. P., Joshi, S. H., Redlich, R., Dannlowski, U., Jorgensen, A., Abbott, C. C., Oltedal, L., & Narr, K. L. (2021). Accounting for symptom heterogeneity can improve neuroimaging models of antidepressant response after electroconvulsive therapy. *Human Brain Mapping*, *42*(16), 5322–5333. https://doi.org/10.1002/hbm.25620

Wang, D., Xue, S.-W., Tan, Z., Wang, Y., Lian, Z., & Sun, Y. (2019). Altered hypothalamic functional connectivity patterns in major depressive disorder. *Neuroreport*, *30*(16), 1115–1120. https://doi.org/10.1097/WNR.0000000000001335

Wang, Y., Gong, N., & Fu, C. (2021). Major depression disorder diagnosis and analysis based on structural magnetic resonance imaging and deep learning. *Journal of Integrative Neuroscience*, *20*(4), 977–984. https://doi.org/10.31083/j.jin2004098

Wolff, J., Schindler, S., Lucas, C., Binninger, A.-S., Weinrich, L., Schreiber, J., Hegerl, U., Möller, H. E., Leitzke, M., Geyer, S., & Schönknecht, P. (2018). A semi-automated algorithm for hypothalamus volumetry in 3 Tesla magnetic resonance images. *Psychiatry Research: Neuroimaging*, *277*, 45–51. https://doi.org/10.1016/j.pscychresns.2018.04.007

Xin, J., Zhang, Y., Tang, Y., & Yang, Y. (2019). Brain Differences Between Men and Women: Evidence From Deep Learning. *Frontiers in Neuroscience*, *13*, 185. https://doi.org/10.3389/fnins.2019.00185

Yun, J.-Y., & Kim, Y.-K. (2021). Graph theory approach for the structural-functional

brain connectome of depression. *Progress in Neuro-Psychopharmacology &*

Biological Psychiatry, *111*, 110401.

https://doi.org/10.1016/j.pnpbp.2021.110401

Zhao, S., Xia, Y., Huang, Y., Zou, H., Wang, X., Chen, Z., Zhou, H., Han, Y., Tang, H.,

Yan, R., Yao, Z., & Lu, Q. (2021). The Correlation Between Thyroid Function,

Frontal Gray Matter, and Executive Function in Patients With Major Depressive

Disorder. *Frontiers in Endocrinology*, *12*, 779693.

https://doi.org/10.3389/fendo.2021.779693

Zhao, Y., Dahmani, L., Li, M., Hu, Y., Ren, J., Lui, S., Wang, D., Kuang, W., Gong, Q.,

& Liu, H. (2022). Individualized functional connectome identified replicable

biomarkers for dysphoric symptoms in first-episode medication-naïve patients

with major depressive disorder. *Biological Psychiatry: Cognitive Neuroscience*

and Neuroimaging. https://doi.org/10.1016/j.bpsc.2021.12.010

Zhou, R., Chen, J., Zhao, G., Wang, Z., Peng, D., Xia, W., Mao, R., Xu, J., Wang, F.,

Zhang, C., Wang, Y., Yuan, C., Su, Y., Huang, J., Yang, T., Wang, C., Cui, L.,

Wang, J., Palaniyappan, L., & Fang, Y. (2021). Neural biomarker of functional

disability in major depressive disorder: A structural neuroimaging study.

Progress in Neuro-Psychopharmacology & Biological Psychiatry, *111*, 110337.

https://doi.org/10.1016/j.pnpbp.2021.110337

Zuo, Z., Ran, S., Wang, Y., Li, C., Han, Q., Tang, Q., Qu, W., & Li, H. (2019).

Asymmetry in cortical thickness and subcortical volume in treatment-naïve

major depressive disorder. *NeuroImage : Clinical*, *21*.

https://doi.org/10.1016/j.nicl.2018.101614