The Canadian Biomarker Integration Network in Depression (CAN-BIND): magnetic resonance imaging protocols

Adolescent depression and brain development: evidence from voxel-based morphometry
The microbiome and mental health: Hope or hype?  
V.H. Taylor

The Canadian Biomarker Integration Network in Depression (CAN-BIND): magnetic resonance imaging protocols  

Adolescent depression and brain development: evidence from voxel-based morphometry  
J. Straub, R. Brown, K. Malejko, M. Bonenberger, G. Grön, P.L. Plener, B. Abler

An artificial neural network model for clinical score prediction in Alzheimer disease using structural neuroimaging measures  
N. Bhagwat, J. Pipitone, A.N. Voineskos, M.M. Chakravarty; Alzheimer’s Disease Neuroimaging Initiative

Changes of motor cortical excitability and response inhibition in patients with obsessive–compulsive disorder  

Randomized controlled trial of a gluten-free diet in patients with schizophrenia positive for antigliadin antibodies (AGA IgG): a pilot feasibility study  

Neural response to emotional faces in monozygotic twins: association with familial risk of affective disorders  
I. Meluken, N. Ottesen, C. Harmer, J. Macoveanu, H. Siebnner, L. Kessing, M. Vinberg, K. Miskowiak

Editors-in-Chief  
Patricia Boksa, PhD  
Ridha Joober, MD, PhD

Editor Emeritus  
Yvon D. Lapierre, MD

Associate Editors  
Paul Albert, PhD  
Martin Alda, MD  
M. Mallar Chakravarty, PhD  
Marco Leyton, PhD  
Glenda M. MacQueen, MD, PhD

Associate Editors (Statistics)  
Aurelie Labbe, PhD  
Norbert Schmitz, PhD

Editorial Board competing interest statements are available at jpn.ca

Editorial Advisory Board  
Galila Agam, PhD, Israel  
Clare Beasley, PhD, Canada  
Stefan Borgwardt, MD, PhD, Switzerland  
Jan Buitelaar, MD, PhD, Netherlands  
Udo Dannlowski, MD, PhD, Germany  
Ronald Duman, PhD, USA  
Klaus P. Ebmeier, MD, United Kingdom  
Stefan Ehrlich MD, PhD, Germany  
Stan Floresco, PhD, Canada  
Guido Frank, MD, USA  
Thomas S. Frodl, MD, Ireland  
Shigenobu Kanba, MD, PhD, Japan  
MATCHER Keshavan, MD, USA  
Naguib Mechawar, PhD, Canada  
Igor Nenadic, MD, Germany  
Patricia Ohrmann, MD, Germany  
Alessandro Serretti, MD, Italy  
Catherine Winstanley, PhD, Canada

Submit your manuscript at  
http://mc.manuscriptcentral.com/jpn

© 2019 Joule Inc.

ISSN 1180-4882
Online Research Papers

E1 Impact of white matter hyperintensity location on depressive symptoms in memory-clinic patients: a lesion–symptom mapping study
A.E. Leeuwis, N.A. Weaver, J.M. Biesbroek, L.G. Exalto, H.J. Kuijf, A.M. Hooghiemstra, N.D. Prins, P. Scheltens, F. Barkhof, W.M. van der Flier, G.J. Biessels; on behalf of the TRACE-VCI study group

E11 Altered white matter connectivity in young people exposed to childhood abuse: a tract-based spatial statistics (TBSS) and tractography study
L. Lim, H. Hart, H. Howells, M.A. Mehta, A. Simmons, K. Mirza, K. Rubia

260 Service Information

Psychopharmacology for the Clinician

287 Management of sexual adverse effects induced by atypical antipsychotic medication
L. Downing, D.D. Kim, R.M. Procyshyn, P. Tibbo
The microbiome and mental health: Hope or hype?

Valerie H. Taylor, MD, PhD

The biggest predictor of future behaviour is past behaviour, but in science we at times seem remarkably unable to learn from past experiences. Take the furor regarding the gut microbiome, especially as it pertains to mental illness. Less than 10 years ago the concept of microbial manipulation to treat mental illness was primarily hypothetical, yet we now have clinical trials underway to look at the viability of fecal microbiota transplantation (FMT) as a treatment for major mental illness. While it is essential that the compelling preclinical work supporting the microbiome gut–brain axis be translated in a clinical population, and indeed that is a ubiquitous recommendation of many preclinical animal studies, it is essential to reflect on just how novel and revolutionary the concept of the gut–brain link really is. We also need to identify lessons to be learned from the past to ensure we recommend science and not science fiction going forward.

Prior to this recent renaissance, the concept of the gut–brain connection was prominent in the early 20th century when terms such as “autointoxication” and “intestinal toxemia” described a process whereby intestinally derived toxins were purported to influence health — often mental health specifically — via internal “poisoning” from the toxic contents of the colon. A paper published in The Journal of the American Medical Association in 1898, for example, suggested that microbial growth in the intestines associated with conditions such as a lack of stomach acid, could, under normal circumstances, be addressed via normal body processes, but the concern was that these detoxification pathways might be overrun in melancholia. This concept of vulnerability toward mental illness creating an environment where relatively normal amounts of stress suddenly result in illness is in keeping with the concept of allostatic load, but lacking sophisticated ways to understand the stress diathesis model and the role of neurotransmitters and inflammation, treatment focused on the identified problem area: the gut. This not only led to treatments targeted toward changing the gut via oral consumption of products comprising things such as lactic acid–producing bacteria, but also involved extreme measures such as gastrointestinal surgery for schizophrenia — a procedure associated with high mortality and little else.

The flawed translation from scientific theory to clinical utility marked the beginning of the end for this school of thought and should be reflected on as the field, buoyed by strong preclinical work, again attempts to transition into the realm of treatment. One reason for the initial departure from the concept of autointoxication was purported to have been that, in parallel to the legitimate scientific interest in the effects of intestinal bacteria on health, charlatans were alert to the financial possibilities offered by the idea that cleaning out the colon could instantly improve well-being and began to sell a variety of alternative health products linked to improving your gut microflora and your sense of well-being. The irony should not be lost, given that the current global wellness industry had a $4.2 trillion market in 2018, and class action lawsuits have been settled for claims that probiotic-enriched foods can “boost the body’s defenses” and “enhance the immune system.”

Thankfully, the concept of gastrointestinal surgery to treat mental illness is no longer accepted, but the new methods being used to manipulate the gut microflora are not without risk. The primary tenet of FMT is that dysbiosis within the human host gut microbiome predisposes an individual to disease. The exact mechanisms through which this occurs have not yet been established, but several potential direct and indirect pathways exist through which the gut microbiota can modulate the gut–brain axis. These pathways include endocrine (cortisol), immune (cytokines) and neural (vagus and enteric nervous system) pathways, and the assumption is that introducing microflora from a healthy individual will help recolonize the system with a microbial pattern more in keeping with wellness either by establishing the new healthy microbiota or by allowing the host to “reset” their own microflora to a pre-illness state. As of yet, there is no clearly defined gold standard profile associated with euthymia, and factors such as diet, smoking and age can affect the microbiome. There is, however, work indicating that certain bacteria...
FOQUEST®
METHYLPHENIDATE HYDROCHLORIDE CONTROLLED RELEASE CAPSULES

Pharmacokinetics

Rapid onset* of action: within 1 hour

Long-acting* duration of action: 16 hours

WIDE RANGE OF STRENGTHS TO ASSIST WITH TITRATION†

- Maximum daily dose in children and adolescents (6 to 17 years) is 70 mg
- Maximum daily dose in adults (≥18 years) is 100 mg

PAEDIATRIC STRENGTHS

25 mg  45 mg  55 mg  70 mg  85 mg  100 mg

† Please see the Product Monograph for complete dosing information.

FOQUEST® (methylphenidate hydrochloride controlled release capsules) is indicated for the treatment of Attention Deficit Hyperactivity Disorder (ADHD) in patients ≥6 years of age.

* Clinical significance has not been established.

See prescribing summary on page 221
Please consult the product monograph at http://purdue.ca/en/products/Foquest-PM for important information relating to:

- Contraindications in persons with known hypersensitivity or idiosyncrasy to sympathomimetic amines; thyrotoxicosis; advanced arteriosclerosis; symptomatic cardiovascular disease; moderate to severe hypertension; glaucoma; history of drug abuse; during or within 14 days following administration of monoamine oxidase inhibitors
- The most serious warning and precaution regarding drug dependence/tolerance
- Other relevant warnings and precautions regarding: long-term effects of methylphenidate; use in patients who are involved in strenuous exercise or activities, use other stimulants, or have a history of sudden/cardiac death; risk of sudden death, stroke and myocardial infarction; use of CNS stimulants in patients with cardiovascular or cerebrovascular conditions; use in patients with hypertension, taking alcohol; increase in seizure frequency, onset or exacerbation of motor or verbal tics; impairment in ability to operate machinery or vehicles; visual disturbances; psychiatric effects, including not for treatment of depression, for or use in treatment or prevention of normal fatigue states; exacerbation of psychosis symptoms in patients with pre-existing psychotic disorder; need to screen for risk of bipolar disorder in patients with comorbid depressive symptoms; need to monitor patients for: signs of suicide-related behaviour; new psychotic or manic episodes, aggressive behaviour, marked anxiety, or agitation; caution regarding concomitant use of methylphenidate and serotonergic drugs (not recommended); precautions regarding priapism, peripheral vasculopathy, including Raynaud’s phenomenon; use in pregnancy or breastfeeding; periodic laboratory testing during prolonged therapy; potential for misuse and dependence
- Conditions of clinical use, adverse reactions, drug interactions, and dosing information which have not been discussed in this piece.

The product monograph is also available by calling us at 1-800-387-5349.

community-based samples, however, which can be interpreted as representing work focused on participants who at baseline were too well to show meaningful impact. The results, therefore, can be used to endorse either viewpoint regarding the therapeutic potential of these compounds in the treatment of mental illness — which is not helpful.

As is true in many areas of mental health research, our diagnostic categorizations may be part of the issue, given that using DSM criteria for depression produces more than 200 different symptom profiles, all of which fit the major depressive disorder (MDD) diagnostic category. It may be, then, that probiotics are helpful for some, but not all, types of MDD. This area of research has the additional challenge of not only needing to identify what type of depression would be amenable to changes in the gut microflora, but also what type of probiotic should be used as an intervention and what the timing of the treatment should be. Results from a germ-free mouse study showed that central nervous system neurotransmission can be profoundly disturbed by the absence of a normal gut microbiota and that this aberrant neurochemical profile is resistant to restoration of a normal gut flora in later life. We also have systematic reviews of studies on the fecal microbiota in anorexia nervosa, the goal being to facilitate weight gain, and a review of the microbiome in the management of obesity, the goal being to facilitate weight loss, which speaks to our need to better understand what the science is and how we are going to translate this knowledge.

There is still a lack of understanding of what we are trying to change with treatment, as there is no microbial profile clearly associated with wellness, or vice versa, from a behavioural perspective. There is growing capacity, however, to move beyond causal associations and to begin to understand from a more mechanistic perspective exactly what the microbiome looks like when linked to wellness and to begin to think of the microbiome as perhaps contributing to the armamentarium of potential biomarkers we have available to help guide treatment. A recent large Dutch study has begun work at a population level to help create a dedicated reference database that will allow for a more sophisticated study of microbial neuroactive potential and provides population-scale evidence for microbiome links to mental health, while also acknowledging the issue of confounders. While provocative, the study also illustrates the need for multicentre work in this area, using validated instruments and solid metagenomic techniques to attempt to apply the cardiology model of thousands and validated instruments and solid metagenomic techniques impart rigour. If this is to be a game changer for mental illness, we need to change our approach. Otherwise we will, again, fail to truly know if a strong biological plausibility holds scientific credibility.

Affiliations: From the Department of Psychiatry, Hotchkiss Brain Institute, University of Calgary, Calgary, Alta., Canada.

Competing interests: V. H. Taylor has grant funding from the Stanley Medical Research Institute and the Weston Family Microbiome Initiative for microbiome research.

References

3. Brower DR. Auto-intoxication in its relations to the diseases of the nervous system. JAMA 1899;575-7.
The Canadian Biomarker Integration Network in Depression (CAN-BIND): magnetic resonance imaging protocols

Glenda M. MacQueen, MD, PhD*; Stefanie Hassel, PhD*; Stephen R. Arnott, PhD; Jean Addington, PhD; Christopher R. Bowie, PhD; Signe L. Bray, PhD; Andrew D. Davis, PhD; Jonathan Downar, MD, PhD; Jane A. Foster, PhD; Benicio N. Frey, MD, PhD; Benjamin I. Goldstein, MD, PhD; Geoffrey B. Hall, PhD; Kate L. Harkness, PhD; Jacqueline Harris, MSc; Raymond W. Lam, MD; Catherine Lebel, PhD; Roumen Milev, MD, PhD; Daniel J. Müller, MD, PhD; Sagar V. Parikh, MD; Sakina Rizvi, PhD; Susan Rotzinger, PhD; Gulshan B. Sharma, PhD, MBA; Claudio N. Soares MD, PhD, MBA; Gustavo Turecki, MD, PhD; Fidel Vila-Rodriguez, MD; Joanna Yu, PhD; Mojdeh Zamyadi, MSc; Stephen C. Strother, PhD; Sidney H. Kennedy, MD; on behalf of the CAN-BIND Investigator Team

Studies of clinical populations that combine MRI data generated at multiple sites are increasingly common. The Canadian Biomarker Integration Network in Depression (CAN-BIND; www.canbind.ca) is a national depression research program that includes multimodal neuro-imaging collected at several sites across Canada. The purpose of the current paper is to provide detailed information on the imaging protocols used in a number of CAN-BIND studies. The CAN-BIND program implemented a series of platform-specific MRI protocols, including a suite of prescribed structural and functional MRI sequences supported by real-time monitoring for adherence and quality control. The imaging data are retained in an established informatics and databasing platform. Approximately 1300 participants are being recruited, including almost 1000 with depression. These include participants treated with antidepressant medications, transcranial magnetic stimulation, cognitive behavioural therapy and cognitive remediation therapy. Our ability to analyze the large number of imaging variables available may be limited by the sample size of the substudies. The CAN-BIND program includes a multimodal imaging database supported by extensive clinical, demographic, neuropsychological and biological data from people with major depression. It is a resource for Canadian investigators who are interested in understanding whether aspects of neuroimaging — alone or in combination with other variables — can predict the outcomes of various treatment modalities.

Introduction

Treatment of major depressive disorder (MDD) is evidence-based, but treatment selection is not personalized to the features of an individual’s illness.1 The discovery of biomarkers — or predictors — of treatment response is a priority in MDD research.2 A major challenge for identifying patient characteristics that predict treatment response is that MDD is a complex, heterogeneous condition. Current diagnostic systems codify depressive symptoms as criteria for MDD,3 but these symptoms are not unique to depression and, even if clustered together, they may not represent a single underlying disease process or treatment substrate.

A growing number of clinical studies are using MRI in an attempt to identify biomarkers of disease (for example, Jack and colleagues4), including depression (see Fonseka and colleagues5 for a recent review of studies using MRI to define markers of outcome in MDD). One approach to the detection of imaging biomarkers is to integrate data from large numbers of patients collected in independent studies. Keshavan and colleagues6 examined the circumstances under which a study could forgo efforts at protocol harmonization and
phantom-based correction, relying only on the power of the data. They performed a scan–rescan study on 20 scanners with similar but nonidentical imaging parameters and determined that, in the absence of protocol harmonization, the sample size required could be in the thousands. The Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium is a collaborative network of researchers who have integrated primarily structural data from more than 12,000 participants and 70 institutions around the world.7 The ENIGMA consortium has a working group focused on MDD that has reported on both subcortical and cortical brain structures.9 However, despite the power of this approach to examine factors such as age of onset and recurrence, ENIGMA’s psychiatric cohorts vary in terms of inclusion and exclusion criteria, duration of illness, the absence or presence of comorbid conditions, treatment history, ethnicity and other factors, limiting investigators’ ability to examine imaging data in the context of relevant clinical variables.5

An alternative approach to combining data from multiple independent studies is to conduct coordinated, multisite imaging studies. Several consortia have established guidelines and protocols for such studies, including the Function Biomedical Informatics Research Network (fBIRN),10 the Alzheimer’s Disease Neuroimaging Initiative (ADNI),4,11 the Mind Clinical Imaging Consortium (MCIC),12 the North American Imaging in Multiple Sclerosis (NAIMS) Cooperator13 and the Ontario Neurodegenerative Disease Research Initiative (ONDRI).14 However, only a few studies to date have employed multimodal, multisite imaging analyses to predict treatment outcomes in MDD.

The international Study to Predict Optimized Treatment in Depression (iSPOT)15 enrolled more than 2000 patients with MDD across 20 sites, but they recruited only 10% of the participants into the neuroimaging substudy, which was conducted at 2 sites,15,16 The iSPOT neuroimaging protocol included high-resolution 3-dimensional T1-weighted scans; diffusion tensor imaging (DTI); and T2-weighted proton density scans, as well as task-based functional MRI (fMRI) sequences to assess cognitive and emotional processing.17 The Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC) study17 enrolled 309 patients with early-onset MDD across 6 sites. The EMBARC neuroimaging protocol included 3-dimensional T2-weighted scans, DTI, arterial spin labelling and task-based fMRI sequences to assess the processing of reward and emotional conflict.

The Canadian Biomarker Integration Network in Depression (CAN-BIND; www.canbind.ca; see Kennedy and colleagues19 for a detailed description). The study included MRI at baseline and after 2 and 8 weeks of treatment. It recruited participants from 6 sites in Canada (ClinicalTrials.gov identifier NCT01655706).

The CAN-BIND-2 study (Canadian rTMS Treatment and Biomarker Network in Depression; CARTBIND) explored the use of repetitive transcranial magnetic stimulation (rTMS), a noninvasive brain stimulation technique approved as a treatment for MDD. The CARTBIND trial is a 3-site study that uses 6 weeks of left dorsolateral prefrontal cortex intermittent theta-burst rTMS in patients with MDD, with the aim of identifying biomarkers of response to rTMS treatment. Scans have been obtained for 205 patients at baseline and within 1 week of completing rTMS therapy (ClinicalTrials.gov identifier NCT02729792).

The CAN-BIND-3 study (Canadian Psychiatric Risk and Outcome Study; PROCAN) is a 2-site study with the goal of improving the ability to identify youth at risk of serious mental illness, including MDD.24 In this study, 240 youth have been recruited, aged 12 to 25 years and at various levels of risk as defined in clinical staging models (e.g., genetic risk only, mild and/or attenuated symptoms, more pronounced but subthreshold symptoms). Participants are scanned at baseline and at 1- and 2-year follow-up, or when symptoms worsen.

The CAN-BIND-4 (Stress and Reward Anhedonia; SARA) single-site study aims to examine stress reactivity and reward responsivity as correlated domains of functioning in depression in 200 participants (100 patients with MDD, 100 healthy controls). Structural and functional brain imaging is being obtained at baseline and 6-month follow-up.

The CAN-BIND-5 (Biomarkers of Suicidality) single-site study has the goal of identifying an integrated biological marker model to predict risk of suicide attempt in MDD, and to test the stability of this model over time. Ninety patients with MDD with and without a history of suicide attempt, as well as 30 healthy controls, are being scanned at a baseline visit and at 1-year follow-up (ClinicalTrials.gov identifier NCT02811998).

The CAN-BIND-9 (Remote Cognitive Remediation for Depression; ReCoRD) single-site study aims to assess the effectiveness of cognitive remediation therapy in 75 participants with MDD who complete computer treatment modules from their homes. Participants are scanned at baseline and after online cognitive remediation, at 12- and 24-week follow-up.

The CAN-BIND-10 (Concussion and Depression Study) single-site study aims to characterize the biological profile of people with mild traumatic brain injury and depression, and
to identify factors that may predict risk of depression after injury. Overall, 100 patients and 25 healthy controls are being scanned at entry into the study.

CAN-BIND participants

Participants are being recruited at 7 Canadian clinical centres: the University Health Network, the Centre for Addiction and Mental Health and Sunnybrook Health Sciences Centre in Toronto, Ontario; St. Joseph’s Healthcare in Hamilton, Ontario; Providence Care Hospital in Kingston, Ontario; Djavad Mowafaghian Centre for Brain Health in Vancouver, British Columbia; and the Mathison Centre for Mental Health Research and Education in the Hotchkiss Brain Institute, Calgary, Alberta. Each site has entered a standardized participation agreement with the Ontario Brain Institute to facilitate the transfer of both raw and processed/deidentified data, in accordance with the Ontario Brain Institute’s governance policy (www.braincode.ca/content/governance) and with any specific conditions required by each institution’s local legislative and ethical policies.

For all studies except CAN-BIND-3 and CAN-BIND-10, patients have a primary diagnosis of MDD, based on structured clinical interview. The CAN-BIND-324 study includes youth aged 12 to 15 years at varying degrees of risk for serious mental illness as defined by a clinical staging model.25 The CAN-BIND-10 study is recruiting patients with traumatic brain injury only and patients with both traumatic brain injury and MDD.

Across all studies, healthy participants for comparison have no history of psychiatric illness or current psychiatric illness as assessed by structured interview. Both patients and healthy participants are excluded if they have an estimated IQ of less than 70 based on the North American Adult Reading Test26; neurologic disease; a history of skull fracture or a severe or disabling medical condition; or a contraindication for MRI. Complete inclusion and exclusion criteria are specific to the various substudies.

CAN-BIND imaging protocols

The CAN-BIND program includes multiple longitudinal studies that employ common neuroimaging elements. Some use additional tasks and modalities, as indicated by the nature of the study. For the main characteristics and protocols for each CAN-BIND study, see Table 1, Table 2, and Appendix 1, Table S1 and Table S2. available at jpn.ca/180036.

The CAN-BIND protocols include the following imaging sequences: a high-resolution 3-dimensional isotropic T1-weighted scan to assess fine anatomical detail and map cortical thickness; DTI to assess microstructural

### Table 1: Overview of CAN-BIND studies highlighting common, standardized data elements

<table>
<thead>
<tr>
<th>CAN-BIND study*</th>
<th>CAN-BIND-1†</th>
<th>CAN-BIND-2</th>
<th>CAN-BIND-3</th>
<th>CAN-BIND-4</th>
<th>CAN-BIND-5</th>
<th>CAN-BIND-9</th>
<th>CAN-BIND-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Drug</td>
<td>rTMS</td>
<td>At-risk youth</td>
<td>Stress and reward</td>
<td>Suicide markers</td>
<td>Cognitive remediation</td>
<td>TBI</td>
</tr>
<tr>
<td>Patient-specific information, diagnosis</td>
<td>MDD</td>
<td>MDD</td>
<td>Youth at risk for severe mental illness; family high risk</td>
<td>MDD</td>
<td>MDD, MDD with suicidal ideation or attempt</td>
<td>MDD, MDD + TBI</td>
<td></td>
</tr>
<tr>
<td>Intervention/treatment</td>
<td>SSRI (escitalopram); aripiprazole</td>
<td>rTMS</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Cognitive remediation</td>
<td>NA</td>
</tr>
<tr>
<td>Patients, n</td>
<td>211</td>
<td>205</td>
<td>200</td>
<td>100</td>
<td>90</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>Controls, n</td>
<td>112</td>
<td>NA</td>
<td>40</td>
<td>100</td>
<td>30</td>
<td>NA</td>
<td>25</td>
</tr>
<tr>
<td>Number of times scanned</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>T1 structural*</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Diffusion tensor imaging</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Resting-state fMRI</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Go/no-go task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Incentive delay task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Working memory task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Breath-holding challenge / breath-hold task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Shifted attention emotion appraisal test</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Probabilistic reward task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Prediction error task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Social cognition task</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
<td>♦</td>
</tr>
<tr>
<td>Other</td>
<td>Face categorization task</td>
<td>Arterial spin labelling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CAN-BIND = Canadian Biomarker Integration Network in Depression; MDD = major depressive disorder; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; SSRI = selective serotonin reuptake inhibitor; TBI = traumatic brain injury.

*Overall, the 7 studies are projected to include approximately 980 patients and 305 controls, for a total of approximately 3000 T1 scans.

†Approximately 600 patient T1 scans, approximately 300 control T1 scans.
### Table 2: Detailed scan acquisition parameters for structural MRI sequences (part 1 of 2)

<table>
<thead>
<tr>
<th>CAN-BIND site</th>
<th>CAN-BIND project</th>
<th>Scanner model</th>
<th>Software version</th>
<th>Coil</th>
<th>T₁-weighted scan, sagittal acquisition</th>
<th>Diffusion tensor imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto Western/ University</td>
<td>CAN-BIND-1</td>
<td>GE 3.0 T</td>
<td>HD16.0_</td>
<td>GE 8HRBRAIN</td>
<td>Repetition time, ms</td>
<td>Repetition time, ms</td>
</tr>
<tr>
<td>Toronto General Hospital</td>
<td>CAN-BIND-2</td>
<td>Signa HDxt</td>
<td>V02_1131.a</td>
<td>GE 8HRBRAIN</td>
<td>Echo time, ms</td>
<td>Echo time, ms</td>
</tr>
<tr>
<td>Centre for Addiction and Mental</td>
<td>CAN-BIND-1</td>
<td>GE 3.0 T</td>
<td>DV24.0_</td>
<td>GE 32Ch Head/GE HNS</td>
<td>Inversion time, ms</td>
<td>Inversion time, ms</td>
</tr>
<tr>
<td>Health</td>
<td>CAN-BIND-2</td>
<td>Discovery MR750</td>
<td>R01_1344.a</td>
<td>Head</td>
<td>Flip angle, degrees</td>
<td>Flip angle, degrees</td>
</tr>
<tr>
<td>McMaster University</td>
<td>CAN-BIND-1</td>
<td>Discovery MR750</td>
<td></td>
<td></td>
<td>Pixel bandwidth</td>
<td>Pixel bandwidth</td>
</tr>
<tr>
<td>University of Calgary</td>
<td>CAN-BIND-3</td>
<td>Phillips 3.0 T</td>
<td></td>
<td></td>
<td>Matrix dimension, pixels</td>
<td>Matrix dimension, pixels</td>
</tr>
<tr>
<td>University of British Columbia</td>
<td>CAN-BIND-2</td>
<td>Intera</td>
<td>DV25.0_R02_1549.b</td>
<td></td>
<td>Voxel dimension, mm</td>
<td>Voxel dimension, mm</td>
</tr>
<tr>
<td>Sunnybrook Health Sciences Centre</td>
<td>CAN-BIND-3</td>
<td>Siemens 3.0 T</td>
<td>DV25.0_R02_1549.b</td>
<td>GE HNS Head</td>
<td>Slices, n</td>
<td>Slices, n</td>
</tr>
<tr>
<td>Queen's University</td>
<td>CAN-BIND-1</td>
<td>Siemens 3.0 T</td>
<td></td>
<td></td>
<td>Acquisition times, min</td>
<td>Acquisition times, min</td>
</tr>
<tr>
<td>Saint Michael's Hospital</td>
<td>CAN-BIND-10</td>
<td>Syrya</td>
<td></td>
<td></td>
<td>Diffusion directions, n</td>
<td>Diffusion directions, n</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Diffusion b value</th>
<th>Diffusion images with</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>b = 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Acquisition times, min</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Diffusion images with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>b = 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Acquisition times, min</td>
</tr>
</tbody>
</table>

**Notes:**
- CAN-BIND-1, CAN-BIND-2, CAN-BIND-3, CAN-BIND-4, CAN-BIND-5, CAN-BIND-6, CAN-BIND-7, CAN-BIND-8, CAN-BIND-9, CAN-BIND-10
- **T₁-weighted scan, sagittal acquisition**
  - Repetition time, ms
  - Echo time, ms
  - Inversion time, ms
  - Flip angle, degrees
  - Pixel bandwidth
  - Matrix dimension, pixels
  - Voxel dimension, mm
  - Slices, n
  - Acquisition times, min

**Diffusion tensor imaging**
- Repetition time, ms
- Echo time, ms
- Flip angle, degrees
- Pixel bandwidth
- Matrix dimension, pixels
- Voxel dimension, mm
- Diffusion directions, n
- Diffusion b value
- Diffusion images with b = 0
- Acquisition times, min
Table 2: Detailed scan acquisition parameters for structural MRI sequences (part 2 of 2)

<table>
<thead>
<tr>
<th>CAN-BIND site</th>
<th>Toronto Western/Toronto General Hospital</th>
<th>Centre for Addiction and Mental Health</th>
<th>McMaster University</th>
<th>University of Calgary</th>
<th>University of British Columbia</th>
<th>Sunnybrook Health Sciences Centre</th>
<th>Queen’s University</th>
<th>Saint Michael’s Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T₂-weighted proton density scan, axial acquisition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition time, ms</td>
<td>6583&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5724&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6004&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6004&lt;sup&gt;d&lt;/sup&gt;</td>
<td>5500</td>
<td>5428</td>
<td>11900&lt;sup&gt;e&lt;/sup&gt;</td>
<td>NA</td>
</tr>
<tr>
<td>Echo time 1, ms</td>
<td>7.3</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>16.8</td>
<td>16.8</td>
<td>8.1</td>
<td>NA</td>
</tr>
<tr>
<td>Echo time 2, ms</td>
<td>87.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
<td>88.0</td>
<td>88.0</td>
<td>105.0</td>
<td>NA</td>
</tr>
<tr>
<td>Flip angle, degrees</td>
<td>90</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>NA</td>
</tr>
<tr>
<td>Pixel bandwidth</td>
<td>326</td>
<td>326</td>
<td>326</td>
<td>326</td>
<td>188</td>
<td>188</td>
<td>326</td>
<td>NA</td>
</tr>
<tr>
<td>Matrix dimension, pixels</td>
<td>192 x 192</td>
<td>192 x 192</td>
<td>192 x 192</td>
<td>192 x 192</td>
<td>192 x 192</td>
<td>192 x 138</td>
<td>192 x 168</td>
<td>NA</td>
</tr>
<tr>
<td>Voxel dimension, mm</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>1.25 x 1.25 x 2.5</td>
<td>NA</td>
</tr>
<tr>
<td>Slices, n</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>NA</td>
</tr>
<tr>
<td>Acquisition times, min</td>
<td>01:47</td>
<td>01:32</td>
<td>01:27</td>
<td>01:30</td>
<td>04:35</td>
<td>02:23</td>
<td>01:47</td>
<td>NA</td>
</tr>
</tbody>
</table>

CAN-BIND = Canadian Biomarker Integration Network in Depression; NA = not applicable.

<sup>a</sup>For n = 42, the repetition time for GE Signa was 7.2 ms.
<sup>b</sup>For n = 59, the repetition time for GE Discovery ranged from 7.2 ms to 7.7 ms.
<sup>c</sup>For n = 11, the repetition time for Siemens was 900 ms.
<sup>d</sup>For n = 42, the echo time for GE Signa was 2.7 ms.
<sup>e</sup>For n = 59, the echo time for GE Discovery ranged from 2.7 ms to 2.9 ms.
<sup>f</sup>For n = 27, the echo time for Philips Achieva was 3.6 ms.
<sup>g</sup>For n = 11, the echo time for Siemens was 2.7 ms.
<sup>h</sup>For n = 11, the inversion time for Siemens was 900 ms.
<sup>i</sup>For n = 78, the pixel bandwidth for GE was 244.
<sup>j</sup>For n = 23, the pixel bandwidth for GE was 122.
<sup>k</sup>For n = 27, the pixel bandwidth for Philips was 191.
<sup>l</sup>For n = 91, the matrix dimensions for GE were 220 x 220.
<sup>m</sup>For n = 42, the repetition time for GE Discovery was 14000 ms. The repetition time for diffusion tensor imaging was adjusted to be consistent (reduced to 8000 ms for GE scanners and 8000 ms for Philips scanners) early in CAN-BIND-1. Subsequent CAN-BIND studies used the reduced repetition time.
<sup>n</sup>For n = 145, the echo time for diffusion tensor imaging sequences ranged from 78 to 85 ms.
<sup)o</sup>For n = 49, the number of slices for GE Discovery was 176.
<sup>p</sup>For n = 145, the echo time for diffusion tensor imaging sequences ranged from 78 to 85 ms.
<sup>q</sup>For n = 49, the pixel bandwidth for GE was 3906.
<sup>r</sup>For n = 19, the pixel bandwidth for Siemens was 1933.
<sup>s</sup>For n = 27, the pixel bandwidth for GE was 3256.
<sup>t</sup>For n = 11, the pixel bandwidth was 1408.
<sup>u</sup>For n = 19, the voxel dimensions was 0.9 x 0.9 x 2.5.
<sup>v</sup>For n = 27, the voxel dimension was 2.4 x 2.4 x 2.4.
<sup>w</sup>For n = 42, the repetition time was 7167 ms.
<sup>x</sup>For n = 7, the repetition time was 5794 ms.
<sup>y</sup>For n = 23, the repetition time was 5901 ms.
<sup>z</sup>For n = 19, the repetition time was 6064 ms.
<sup>a</sup>For n = 11, the repetition time was 11670 ms.
<sup>b</sup>For n = 42, the echo time was 88.32 ms.
white-matter integrity; and resting-state and task-based blood-oxygenation-level-dependent fMRI sequences to assess functional networks and pathways. The CAN-BIND-3 study also uses arterial spin labelling to measure cerebral blood flow. Protocols have been informed by a review of the relevant literature, consultation with other experts in the field and group consensus, taking into account each scanner’s capabilities.

Six scanner models are used across the clinical sites, mandating extensive and ongoing quality-control processes: a Discovery MR750 3.0 T (GE Healthcare), a Sigma HDxt 3.0 T (GE Healthcare), a MAGNETOM Trio (Siemens Healthcare), a MAGNETOM Skrya (Siemens Healthcare), an Achieva 3.0 T (Philips Healthcare) and an Intera 3.0 T (Philips Healthcare).

Stimulus sizes, instructions to participants and support materials are standardized across sites. All behavioural data are captured using E-Prime version 2.0 Professional (Psychology Software Tools). For CAN-BIND-5 and CAN-BIND-10, PsychoPy, Inquisit (Millisecond) and Presentation (www.neurobs.com/) are also used. Guidelines and practices have been established for instructing participants to remain still throughout the scan, for applying a fiducial marker on the right temple, and for collecting respiratory bellows and peripheral gating (pulse oximetry) data using standard instruments provided by each manufacturer.

Whole-brain T₁-weighted structural scan

Whole-brain T₁-weighted structural scans are noninvasive, readily acquired and, because they are relatively short, generally well tolerated; these are features that may be important for identifying a potential biomarker. Structural MRI studies in patients with MDD have revealed widespread corticocytic differences in grey matter and white matter, suggesting that there are detectable alterations in the structure of key brain regions that could inform clinically relevant outcomes. Studies examining how well structural MRI data may be able to diagnose depression report accuracy rates of 48% to 91%. Some studies have reported that structural alterations predict outcomes of treatment at the group level.

The T₁-weighted scans are acquired with a 3D isotropic resolution of 1 mm. For further detail on whole-brain T₁-weighted imaging parameters, see Table 2. Information to confirm participant orientation is collected by placing a small vitamin E capsule on the right temple as a stereotactic marker. Further information is included in Setup and Quality Assurance of MRI Protocols.

Whole-brain DTI

Diffusion tensor imaging studies have demonstrated altered white-matter microstructural abnormalities in patients with MDD. Decreased fractional anisotropy, a proxy measure of the directionality of diffusion, has been reported in patients with MDD in the frontal and occipital (fusiform) regions. Fibre tracking has revealed the involvement of similar structures in MDD. White-matter alterations have predicted treatment outcomes with up to 65% accuracy. In another study, elevated baseline fractional anisotropy in tracts connecting to the right amygdala has been associated with remission following SSRI treatment.

The CAN-BIND DTI acquisition protocol employs a singleshot, spin-echo, echo planar imaging sequence with diffusion sensitizing gradients applied in 31 noncollinear directions (b = 1000 s/mm²) and 6 volumes with b = 0 s/mm². For CAN-BIND-3, diffusion sensitizing gradients were applied in 45 noncollinear directions, with 8 images collected at b = 1000 s/mm² and 8 images collected at b = 2500 s/mm². Increasing the number of diffusion-encoded directions improves the accuracy and/or robustness of diffusion tensor estimation, and having more directions allows for the removal of any corrupted directions (e.g., due to motion/movement). See Table 2 for further details on the parameters for whole-brain DTI.

Resting-state fMRI

Resting-state fMRI allows for the identification of task-independent and spontaneous neural activation that coincides temporally to form neural networks such as the default mode network (e.g., see Greicius and colleagues), the salience network or cognitive control network (e.g., Menon, Menon and Uddin, or Seeley and colleagues), and the affective network. The default mode network shows abnormal patterns of functional connectivity in MDD that may normalize following treatment or may be associated with treatment resistance. Resting-state data are collected over a 10-minute scan during which participants are instructed to lie still, keep their eyes open and focus on a fixation cross. Standardized instructions are used across sites. Images are obtained using a whole-brain T₁*-sensitive blood-oxygen-level-dependent echo planar imaging series, with a repetition time of 2000 ms, an echo time of 30 ms and voxel dimensions of 4 mm × 4 mm × 4 mm, kept constant across sites and scanners. See Table 3 for further details on the parameters for resting-state fMRI.

Task-based fMRI

Task-based fMRI studies suggest that there may be different patterns of change associated with specific treatments or classes of treatment. The CAN-BIND substudies test treatment- and population-specific questions, using cognitive-functional tasks that are described in detail in Appendix 1. Task-related instructions are standardized and given before the scan sessions. Each site uses a comparable, custom-manufactured, magnet-compatible input device (www.mrn.org/collaborate/imaging-equipment) to record participants’ responses. Acquisition parameters are similar to those for resting-state fMRI, and are listed in detail in Appendix 1, Table S1 and Table S2.

Arterial spin labelling

Arterial spin labelling perfusion MRI measures regional cerebral blood flow and may be used to study subtle brain
perfusion changes in psychiatric illnesses. Perfusion patterns may hold promise as objective biomarkers for tracking illness progression, as well as pharmacological/treatment effects in various neuropsychiatric disorders.77

Data storage

Clinical data are collected and stored in the Ontario Brain Institute’s Centre for Ontario Data Exploration (Brain-CODE; www.braincode.ca/; Vaccarino and colleagues78). This online neuroinformatics platform allows researchers to collaborate across distances and work efficiently at multiple sites. Brain-CODE is deployed at the Centre for Advanced Computing at Queen’s University in Kingston, Ontario. The Centre for Advanced Computing is a member of the Compute Canada high-performance computing consortium, which supports regulatory-compliant processes for securing the privacy of health care data (https://cac.queensu.ca/overview). Online clinical and neuroimaging data are accessed on secure websites via restricted portals that require unique usernames and passwords for each member of the study team. User profiles are assigned only to study personnel who require access to enter and verify data, and credentials for each user are vetted by the program manager.

The SPReD database (originally the Stroke Patient Recovery Research Database) is a comprehensive online repository powered by the open-source Extensible Neuroimaging Archiving Toolkit (XNAT) imaging informatics platform,79,80 where neuroimaging data are uploaded and stored. Structural and functional MRI data are uploaded from each site as Digital Imaging and Communications in Medicine (DICOM) images. Supplementary records, such as behavioural and physiological data, and session notes associated with an imaging session, are uploaded through a special subprocess.

Neuroinformatics framework

The CAN-BIND neuroinformatics framework consists of software, tools, pipelines and procedures designed to ensure high-quality data acquisition, databasing, archiving, assessment, analysis and tracking, an overview of which is shown in Figure 1. The primary platform for this set of tools is XNAT/SPReD, provided through Brain-CODE. In addition to the MRI data being captured and managed through XNAT/SPReD, other study-related data are captured using OpenClinica and RedCap. A visualization “dashboard” built using SpotFire (http://spotfire.tibco.com/) is used to upload aggregated data tracking and analytics results from phantom data (see Fig. 2 and Fig. 3).

CAN-BIND quality control and quality assurance procedures

The importance of quality assurance and control in multisite studies is recognized.81 The full spectrum of data quality control and data quality assurance methods was implemented early in CAN-BIND-1. These methods are described in the sections that follow and have been applied to most of the CAN-BIND substudies. The CAN-BIND-2 and CAN-BIND-3 studies have not been uploading their data to SPReD, so the automated adherence checks described here do not apply to them.

Quality control

Data file-naming convention and adherence checks

Participants are assigned unique identification codes, which are standardized to contain a program code (3 letters), a study number (2 digits), a site identification code (3 letters) and a participant number (4 digits; e.g., CBN01_UCA_0001). These file-naming conventions are applied to MRI and behavioural data files. A pipeline assessing the consistency of naming conventions is implemented in XNAT/SPReD; if noncompliance is detected, notification is sent to relevant study personnel asking them to implement corrections, with

Fig. 1: Overview of the CAN-BIND neuroinformatics framework. Data from each site is uploaded to Brain-CODE (XNAT/SPReD), where specifically designed pipelines check the data for compliance with scan acquisition parameters, naming convention and completeness. Automatic messages are sent to initiate manual QC. The CAN-BIND neuroinformatics framework also includes pipelines for the analysis of phantom data. CAN-BIND = Canadian Biomarker Integration Network in Depression; fBIRN = Functional Biomedical Informatics Research Network; QC = quality control; SPReD = originally named the Stroke Patient Recovery Research Database; XNAT = Extensible Neuroimaging Archiving Toolkit.
follow-up until corrections are performed. The data will not undergo subsequent quality-control checks until file-naming conventions have been adhered to.

Parameter adherence checks of MRI protocols
Also implemented in SPreD is a quality-control pipeline for MRI protocols, which compares the acquisition parameters of newly uploaded scans against a reference protocol. Reference protocols have been established for each site and scanner type, taking into account the fact that scan parameters are necessarily different among scanners and manufacturers. The reference protocol defines the sequences and appropriate acquisition parameters (values) for each sequence. If discrepancies are identified between the data uploaded and the reference protocol, e-mail notifications are sent to study personnel, asking them to identify causes for adherence check failures and pointing to the need for possible rescanning.

Image quality
It is necessary to obtain images of sufficient subjective quality, free of motion artifacts, covering a full field of view and free of other scanner-related artifacts in order to process the data through various pipelines. Certain sequences, such as resting-state fMRI, are more susceptible to motion and other artifacts. Others, such as $T_1$-weighted images, are of such paramount importance that tolerance for motion or other

---

**Fig. 2:** Examples of data quality tracking and assessment pipelines. Phantom data are tracked longitudinally to monitor adherence and data quality of imaging protocols. Illustrated here is an example where spiking in the overall mean signal intensity across acquired images at one data acquisition site (light blue) was tracked to be related to its SNR and its SFNR. (A) Mean signal longitudinal: this metric tracks the average overall signal intensity across all voxels and images, per scanning session. (B) SNR longitudinal: this metric tracks the average overall SNR. The mean SNR is the static spatial noise $\times$ image across a $21 \times 21$ voxel region of interest centred on the image. The signal summary value is the average of the signal image across this same region of interest. Then, SNR = (signal summary value)/√(variance summary value/number of time points). (C) SFNR longitudinal: the SFNR is the voxel-wise ratio of the temporal variance standard deviation and temporal mean intensity of the 4-dimensional phantom image after quadratic detrending. The SFNR summary value is the mean SFNR value within the evaluation region of interest (a $21 \times 21$ voxel region in the centre of the image). SFNR = signal-to-fluctuation-noise ratio; SNR = signal-to-noise ratio.
artifacts is low because they influence the quality of the data and the usability of other sequences, which are typically coregistered to $T_1$-weighted scans. Trained expert quality-control raters are automatically notified when new data are uploaded to SPReD. They perform visual assessment of the MRI data image quality using the SPReD interface. The quality-control raters have received training via ONDRI, based on the data quality control protocol from the Centre for Brain Science at Harvard University.22 Raters compare their assessments and comments on scan quality for subsets of data collected at participating CAN-BIND sites. Each imaging sequence is reviewed independently for quality, including full-brain coverage (on a 2-point scale: complete or incomplete), motion and other image artifacts (on a 3-point scale: none, mild or severe), based on the data quality-control protocol.22 Imaging that has insufficient coverage, excessive motion as identified by visual inspection of rigid uniform stripes running horizontally across the brain22 or other imaging artifacts that may interfere with future processing and usability are marked as questionable or unusable, depending on severity. If images are flagged as unusable, they are unavailable for subsequent analysis, and a request is made to the study site to rescans the participant whenever feasible. An upload delay dashboard also serves to inform program managers of the delay time in uploading data once it has been acquired.

**Assessment of site differences**

Cross-site $T_1$ piloting included a travelling participant or “human phantom,” who travelled to each CAN-BIND-1 site for anatomic scans to document within and between-site variance.

**Setup and quality assurance of MRI protocols**

**Setup of scan parameters**

Prior to study launch, scan parameters from DICOM header files were examined to match scan parameters

---

**Fig. 3:** Examples of data quality tracking and assessment pipelines. Phantom data are tracked longitudinally to monitor adherence and data quality of imaging protocols. Illustrated here are examples where the mean intensity of ghost-only voxels showed deviations; investigation and explanation of these anomalies are listed in 1, 2 and 3, below. Mean bright ghost longitudinal: ghost metrics are calculated for each volume by taking a dilated mask (“original mask”) of the data, and shifting it by $N/2$ voxels in the appropriate axis to create a “ghost mask.” Whereas the mean intensities of those voxels in the ghost mask and not in the original mask is the “mean ghost” value, the “mean bright ghost” is the mean intensity of the top 10% of ghost-only voxels. (1) Anomaly: investigation led to protocol adjustments. (2) Receiver coil failure: addressing failure resulted in data returning to the level seen previously. (3) Anomalies, investigation: corresponding human functional MRI scans acquired around this date appeared fine; subsequent phantom scans were fine.
Table 3: Detailed scan acquisition parameters for resting-state functional MRI sequences

<table>
<thead>
<tr>
<th>CAN-BIND site</th>
<th>Centre for Addiction and Mental Health</th>
<th>University of Calgary</th>
<th>University of British Columbia</th>
<th>Sunnybrook Health Sciences Centre</th>
<th>Queen’s University</th>
<th>Saint Michael’s Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAN-BIND project</td>
<td>CAN-BIND-1</td>
<td>CAN-BIND-2</td>
<td>CAN-BIND-1</td>
<td>CAN-BIND-3</td>
<td>CAN-BIND-2</td>
<td>CAN-BIND-3</td>
</tr>
<tr>
<td>Scanner model</td>
<td>GE 3.0 T</td>
<td>GE 3.0 T</td>
<td>GE 3.0 T</td>
<td>GE 3.0 T</td>
<td>Phillips 3.0 T</td>
<td>Philips 3.0 T</td>
</tr>
<tr>
<td>Coil</td>
<td>GE 8HRBRAIN</td>
<td>GE 8HRBRAIN</td>
<td>GE 32Ch Head</td>
<td>GE HNS Head</td>
<td>SENSE-Head-8</td>
<td>SENSE-Head-8</td>
</tr>
</tbody>
</table>

Resting-state functional MRI

- Echo time, ms: 30.0, 30.0, 30.0, 30.0, 30.0, 30.0, 30.0, 30.0
- Field of view: 256, 256, 256, 256, 256, 256, 1536 (mosaic), 256
- Flip angle, degrees: 75.00, 75.00, 75.00, 75.00, 75.00, 75.00, 75.00, 30.0
- Matrix dimension, pixels: 64 × 64, 64 × 64, 64 × 64, 64 × 64, 64 × 64, 64 × 64, 64 × 64, 64 × 64
- Voxel dimension, mm: 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4, 4 × 4 × 4
- Volumes, n: 300, 300, 300, 300, 300, 300, 300, 300
- Slices, n: 36, 36, 36, 36, 36, 36, 36, 36
- Acquisition times, min: 10:00, 10:00, 10:00, 10:00, 10:06, 10:06, 10:00, 10:08

Note: CAN-BIND = Canadian Biomarker Integration Network in Depression. For n = 11, echo time = 25.0 ms. For n = 27, flip angle = 90 degrees. For n = 40, number of slices = 40. For n = 7, number of slices = 40. For n = 11, number of slices = 40.
The CAN-BIND MRI protocols
Sunnybrook Health Sciences Centre, the second data acquisition site for CAN-BIND-3.

Setup of fMRI paradigms
To standardize the viewing angle for fMRI task stimuli, a standard grid was displayed at each site, viewing distance was measured, and the visual angle of the projected image was calculated. Consistent cross-site viewing angle was established using specific display parameters in the E-Prime software for each site. Across sites, the version of the E-Prime stimulus display software was matched. Button responses and ASCII key codes were confirmed and used in site-specific E-Prime task versions. Data files produced by each paradigm were examined to confirm that the proper response information was being acquired and logged.

Sites were also provided with a scripted set of instructions to be issued before resting-state scans, as well as a standardized fixation cross for participants to focus on during the resting state scan. A set of participant orientation/training slides were instituted for functional tasks. Randomization schedules were provided for the functional task version administered (e.g., A/B/C for the go/no-go task) and task order between, for example, go/no-go and reward tasks. For detail on MRI tasks, see Appendix 1. Study coordinators were provided with a guide to follow when checking the fidelity of the acquired behavioural data. Finally, conference calls were held with the research coordinators at each site to ensure that standard operating procedures were communicated and instituted.

Discussion
The neuroinformatics procedures and pipelines employed in CAN-BIND address many challenges associated with combining MRI data from multisite studies. Considerable effort has been focused on the image acquisition protocols, and procedures have been implemented — automated, where possible — to ensure the ongoing quality of the images. We recognize, however, that residual differences in neuroimaging data collected across different sites and MRI vendors will likely still exist.

The “reproducibility in science crisis” has required that imaging studies examine common approaches to study design, monitoring and interpretation. Issues underlying the difficulty with replication are multifaceted, and protocols are emerging to ensure that imaging studies are well planned, well executed and well reported. This includes making the details of how studies are designed, executed and analyzed more apparent and transparent. This paper aims to provide methodological detail for the CAN-BIND studies in a transparent and comprehensive manner. As evident from Figure 1, there are common data elements across the CAN-BIND program sub-studies, specifically for 3D anatomic scans, resting-state fMRI and DTI. Scan parameters (as detailed in Table 2 and Table 3) are as comparable and compatible as scanner manufacturer and type allow. Quality-control procedures, such as checking protocol adherence for participant scans and manual quality control of acquired data, are performed for most sub-studies, based on an agreed-upon protocol. For example, although CAN-BIND-2 and CAN-BIND-3 are not currently uploading data to SPReD for automatic protocol adherence checks, data are being manually inspected for data quality.

Limitations
Although we consider it a strength that the CAN-BIND protocol is applied across participants with a wide age range (12 to 70 years), age-related differences will need to be assessed with caution, as will differences in sex and other demographic factors. We did not assess for the presence of cerebrovascular disease in our sample, although there is an association between cardiovascular disease and MDD, but also with MDD and other medical conditions. Given the relatively young age of our samples (e.g., Lam and colleagues, Addington and colleagues, Santesteban-Echarri and colleagues and Kennedy and colleagues), this is unlikely to be a driving factor in neuroimaging results, but medical comorbidity is an important consideration in studies of psychiatric disease. No routine screening for substance use was performed, potentially affecting our findings. As noted above, CAN-BIND-2 and CAN-BIND-3 are not subject to the automatic adherence checks that would result from uploading to SPReD.

Conclusion
The CAN-BIND program is unusual in that it uses a suite of common imaging protocols across a variety of studies that examine predictive markers of response to various treatment modalities in MDD. Although each CAN-BIND substudy is expected to yield valuable information, the consistent protocols, centralized data collection and quality control that will eventually allow for cross-study investigations is likely to be the greatest strength of CAN-BIND. Deidentified CAN-BIND data eventually will be shared by the Ontario Brain Institute with other collaborators and third parties for research purposes. These data sets may inform clinical research teams with similar data sets comparing MDD with other psychiatric conditions, or comparing different treatment modalities. Thus, rigorous, recorded quality control of CAN-BIND neuroimaging and related data are crucial for ensuring the value of this data set to the greatest number of investigators. When fully realized, the CAN-BIND data set will provide a comprehensive resource for researchers interested in predictors, moderators and mediators of response to treatment in MDD.

Acknowledgements: CAN-BIND is an Integrated Discovery Program carried out in partnership with, and with financial support from, the Ontario Brain Institute, an independent nonprofit corporation funded partially by the Ontario government. The opinions, results and conclusions are those of the authors, and no endorsement by the Ontario Brain Institute is intended or should be inferred. Additional funding is provided by the Canadian Institutes of Health Research, Lundbeck, Bristol-Myers Squibb and Servier. Funding and/or in-kind support is also provided by the investigators’ universities and academic institutions. All study medications are independently purchased at wholesale market values.

Affiliations: From the Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, Alta., Canada
Contributors: G. MacQueen, S. Hassel, J. Addington, C. Bowie, S. Bray, J. Downar, J. Foster, B. Frey, B. Goldstein, K. Harkness, C. Lebel, R. Milev, D. Müller, S. Parikh, S. Rizvi, S. Rotzinger, C. Soares, S. Strother and S. Kennedy conceived and designed the article. R. Lam acquired the data. S. Hassel, S. Arnott, A. Davis, J. Harris, G. Sharma, J. Yu, M. Zamyadi and S. Strother analyzed and interpreted the data. G. MacQueen and S. Hassel wrote the article, which all authors revised. Authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

Members of the CAN-BIND Investigator Team: www.canbind.ca/our-team

References


42. Frodl T, Meisenzahl EM, Zill P, et al. Reduced hippocampal volumes associated with the long variant of the serotonin transporter polymorphism in depression. *Arch Gen Psychiatry* 2004;61:177-83.


---

**JPN’s top viewed articles***

1. N-Acetylcysteine in psychiatry: current therapeutic evidence and potential mechanisms of action
   Dean et al.
   *J Psychiatry Neurosci* 2011;36(2):78–86

2. Why is depression more prevalent in women?
   Albert

3. Mental illness is like any other illness: a critical examination of the statement and its impact on patient care and society
   Malla et al.

4. Inappropriate benzodiazepine use in elderly patients and its reduction
   Tannenbaum

5. Mindfulness-based treatments for posttraumatic stress disorder: a review of the treatment literature and neurobiological evidence
   Boyd et al.
   *J Psychiatry Neurosci* 2018;43(1):7–25

6. Treatment of ADHD in patients with bipolar disorder
   Girard and Joobter
   *J Psychiatry Neurosci* 2017;42(6):E11–12

7. On the neurobiology of hallucinations
   Boksa

8. Autism spectrum disorders, schizophrenia and diagnostic confusion
   Woodbury-Smith et al.
   *J Psychiatry Neurosci* 2010;35(5):360

9. Add-on lithium for the treatment of unipolar depression: Too often forgotten?
   Jollant
   *J Psychiatry Neurosci* 2015;40(1):E23–4

*Based on page views on PubMed Central of content published in the last 10 years.
Updated June 11, 2019

---
Adolescent depression and brain development: evidence from voxel-based morphometry

Joana Straub, PhD; Rebecca Brown, PhD; Kathrin Malejko, MD; Martina Bonenberger, PhD; Georg Grön, PhD; Paul L. Plener, MD; Birgit Abler, MD

Background: Investigating adolescents and young adults may provide a unique opportunity to understand developmental aspects of the neurobiology of depression. During adolescence, a considerable physiologic reorganization of both grey and white matter of the brain takes place, and it has been suggested that differences in grey-matter volumes during adolescence may reflect different maturational processes. Methods: We investigated grey-matter volumes in a comparatively large sample (n = 103) of adolescents and young adults (aged 12 to 27 years), 60 of them with a diagnosis of current depression. Results: Replicating previous studies, we found a clear whole-brain effect of age: the older the participants, the lower their global grey-matter volumes, particularly in the paracingulate and prefrontal cortices. Contrasting depressed and healthy youth in a whole-brain approach, we found greater grey-matter volumes in the dorsolateral prefrontal cortex of those with depression. Furthermore, a region-of-interest analysis indicated lower grey-matter volumes in the hippocampus in participants with depression compared with healthy controls. Limitations: The present study was limited because of a skewed sex distribution, its cross-sectional design and the fact that some participants were taking an antidepressant. Conclusion: During adolescence, restructuring of the brain is characterized by marked decreases in prefrontal grey-matter volumes, interpreted as a correlate of brain maturation. Findings of greater volumes in the prefrontal cortex, particularly in younger adolescents with depression, may suggest that these participants were more prone to delayed brain maturation or increased neuroplasticity. This finding may represent a risk factor for depression or constitute an effect of developing depression.
to reach peak volumes earlier in life (around age 7 to 8 in girls and around age 10 in boys), grey-matter volume in frontal and parietal cortices peaks around age 10 to 12 and in the temporal cortex only around age 16, with girls showing earlier peaks than boys. Consequently, areas related to motor or sensory functioning, such as somatosensory or visual cortices, seem to mature earlier than higher-order association areas. Similarly, in terms of cortical thickness (as a measure of another aspect of grey-matter structure) the dorsolateral prefrontal cortex (dLPFC) reaches adult levels particularly late. Differences in grey-matter volumes during adolescence in depressed versus healthy young people may therefore reflect different maturational processes. Influences of age and sex on grey-matter volumes could explain different results for grey-matter volumes in prefrontal areas. There is meta-analysis evidence for volume reductions in adults, and evidence of increased and decreased volumes in different groups of adolescents. However, for most brain regions, it has been suggested that lower grey-matter volumes may reflect greater maturity of the brain during adolescence. Findings of greater prefrontal cortex volumes, particularly in younger depressed adolescents, may suggest that these individuals are more prone to delayed brain maturation as a risk factor for developing depression. Alternatively, increased prefrontal volumes may also be linked to neuroplasticity processes. In this vein, the rumination and compulsive symptoms frequently seen in depression have been suggested to relate to increased prefrontal volumes.

For subcortical regions such as the hippocampus, the situation is different: in healthy adolescents and young adults, hippocampal volumes grow continuously. However, in depressed adolescents (as in adults), decreased volumes have been reported fairly consistently. A possible reason for this may be that maturation in this structure follows other principles, and a decrease in volume is not a sign of maturational processes but of already affected structural integrity as a consequence of depression. A further explanation for lower hippocampal volumes might be that the compulsive brooding commonly seen in depression could interfere with the functioning of episodic memory, which in turn could adversely affect the volume of the hippocampus.

In the current study, we investigated grey-matter volumes in a reasonably large sample of 103 adolescents, all after the onset of puberty, and young adults (age range 12 to 27 years). We hypothesized that including both adolescents and young adults would allow us to replicate earlier findings related to maturational processes in the brain, irrespective of diagnosis. We expected to find significant effects of age in prefrontal and temporal brain regions, where restructuring is assumed to take place during this period of life (with greater grey-matter volumes in younger participants). We expected the somatosensory and visual cortices, as well as phylogenetically older brain areas such as the hippocampus and amygdala, to show no such age effects irrespective of diagnosis, because maturational processes in these areas should be less evident after the onset of puberty. Furthermore, we expected to replicate the findings of 2 previous studies on grey-matter volumes in larger samples of depressed adolescents and youth in addition to the decreased hippocampal volumes found by Jaworska and colleagues, Hagan and colleagues reported effects of depression on grey-matter volume, particularly in the thalamus and ACC.

We formulated 2 main hypotheses. First, using whole-brain analysis, we expected to replicate an effect of age in the entire sample: that cortical volumes (particularly in phylogenetically newer areas of the brain such as the frontal, parietal and temporal cortices) decrease with age, in line with previous studies (hypothesis 1). In a second whole-brain analysis, we investigated the hypothesis that depressed adolescents would have relatively increased prefrontal cortex volumes compared to healthy controls, as has been found inconsistently in smaller samples (hypothesis 2). We expected that our sample — which had a skewed distribution toward younger participants, where maturational effects should be relatively large — would be well suited to detecting such differences.

As a secondary goal, we hypothesized that we would replicate the finding of decreased hippocampal volumes in depressed adolescents relative to controls using a region-of-interest (ROI) analysis, as consistently reported by Jaworska and colleagues and in several studies with smaller sample sizes (hypothesis 3). We also investigated the effects of illness on grey-matter volume in the ACC (ROI analysis) and thalamus (whole-brain analysis) as previously reported by Hagan and colleagues, predicting decreased grey-matter volumes in depressed participants (hypothesis 4).

Methods

Participants

We obtained participants’ morphological and clinical data as part of 2 functional imaging projects on the effects of cognitive-behavioural therapy in depressed adolescents and on depression and nonsuicidal self-injury in adolescents and young adults (part of the project published in Groschwitz et al.). From these studies, we included data from depressed patients (before cognitive behavioural therapy) and healthy controls aged 12 to 27 years. After excluding 1 patient and 1 control because of poor data quality, we had morphological data for 103 participants. All 60 participants in the depression group met DSM-IV criteria for current major depressive disorder at the time of the MRI. All were inpatients or outpatients of the Department of Child and Adolescent Psychiatry and Psychotherapy or of the Department of Psychiatry and Psychotherapy of Ulm University, or they were outpatients from a private child and adolescent psychiatry practice in Ulm, Germany. Exclusion criteria were a current or previous diagnosis of bipolar disorder, schizophrenia or substance abuse; an intellectual disability; or a major somatic or neurologic disorder. Of the patients with depression, 20 were taking psychotropic medication at the time of the scan, usually antidepressants for their current episode. Of these, 9 had a history of current or previous long-term medication of more than 2 months. Four additional patients were currently unmedicated but had a history of previous long-term medication. In the control group, we included only...
participants who had never been diagnosed with any psychiatric disorder and who were matched for age, education and sex. None of the controls had a history of current or previous psychotropic medication. All participants had reached puberty as indicated by Tanner stage 3 in males and menstruation in females. All participants and caregivers (in the case of minors under age 18 years) provided written informed consent. Procedures were carried out in accordance with the Declaration of Helsinki (2013), and the studies were approved by the Institutional Review Board of Ulm University.

Diagnoses were assessed using the German version of the clinical interview Schedule for Affective Disorders and Schizophrenia for School-Age-Children-Present and Lifetime (K-SADS-PL) for DSM diagnoses in adolescents up to age 18 years, or the Structured Clinical Interview for DSM diagnoses (SCID) in young adults. To assess current depressive symptoms, we used the Beck Depression Inventory second edition (BDI-II), German version.

**Structural MRI data acquisition**

We used a 3.0 T Siemens MAGNETOM Allegra Scanner (Siemens) equipped with a head coil to obtain MRI data. We acquired anatomic high-resolution T1-weighted images using a magnetically prepared rapid acquisition gradient echo sequence (MPRAGE: 1 × 1 × 1 mm voxels, band width 130 Hz/pixel, repetition time 2500 ms, inversion time 1.1 s, echo time 4.57 ms, flip angle 12°, field of view 256 × 256, 192 sagittal slices) as part of a larger imaging protocol.

**Statistical analysis of behavioural data**

We performed statistical analyses using the Statistical Package for the Social Sciences 21 (SPSS Inc.). We computed between-group differences by means of 2-sample t tests and \( \chi^2 \) tests. For correlational analyses, we applied the Pearson coefficient. All tests were performed with levels of significance established at \( p < 0.05 \) (2-tailed).

**Statistical analysis of structural MRI data**

**Preprocessing**

We conducted image preprocessing using the Computational Anatomy Toolbox for SPM 12 (CAT12, http://dbm.neuro.uni-jena.de/cat12/) with the following steps: normalization, segmentation and quality check for sample homogeneity. Using standard settings of the toolbox, we normalized data into Montreal Neurological Institute (MNI) space and segmented them into grey matter, white matter and cerebrospinal fluid using the SPM12 tissue probability maps for spatial registration and segmentation. We conducted spatial smoothing as the final step of preprocessing with a Gaussian kernel of 6 mm full width at half maximum using SPM 12 standard routines.

**Whole-brain analyses**

We assessed the effect of age (hypothesis 1: whole-brain analysis) on grey-matter volumes in the whole brain using a simple regression analysis across the entire group of participants, irrespective of diagnosis. We tested whole-brain group differences between healthy controls and patients (hypothesis 2: whole-brain analysis) for significance using the 2-sample t test module for unpaired samples in SPM 12. For both analyses, we included age and sex as covariates, because previous investigations showed a clear influence of these variables on grey-matter volumes. Sex was a nuisance covariate in both analyses; age was the variable of interest for hypothesis 1 and a nuisance covariate for hypothesis 2. We added long-term psychotropic medication (> 2 months currently or in the past; \( n = 13 \)) as another covariate of no interest to the model. Thresholds for both analyses were set at \( p < 0.001 \) at the voxel level, together with a family-wise error (FWE) correction for multiple comparisons at \( p_{\text{FWE}} < 0.05 \) at the cluster level. We extracted estimates of grey-matter volumes from regions with a significant statistical effect to visualize effects.

**Replication of results from previous MRI studies**

To further investigate the validity of our data, we directed additional analyses toward replication of the main results of previous studies from other research groups in similar samples, particularly those by Jaworska and colleagues and Hagan and colleagues. As in Jaworska and colleagues, we tested an a priori hypothesis (hypothesis 3, ROI analysis) for differential effects in the hippocampus using an ROI approach and expecting smaller volumes in depressed youth. Because the tracing procedures used by Jaworska and colleagues to delineate the left and right hippocampus were not available at our site, we used the hippocampus templates from the atlas for automated anatomic labelling accessible via the WFU Pick Atlas for SPM (http://fmri.wfubmc.edu/software/PickAtlas). We thresholded the map at \( p < 0.01 \) at the voxel level. To control for multiple comparisons, we used a cluster-extent threshold of \( p_{\text{FWE}} < 0.05 \) in combination with the small-volume correction (SVC) in SPM, applying the hippocampus masks bilaterally.

To study the differential effects in the thalamus as one main finding of Hagan and colleagues, we first performed an analysis at the whole-brain level, expecting greater thalamic volumes, particularly in younger depressed participants than in controls. The second main finding in that study was a differential effect in the ACC (hypothesis 4, ROI analysis). Like Hagan and colleagues, we used the atlas for automated anatomic labelling accessible via the WFU pickatlas for SPM and combined templates for “Cingulum_Ant_R” and “Cingulum_Ant_L” to delineate a bilateral ACC ROI. At the voxel level, we used the same threshold for the ROI analysis as Hagan and colleagues (\( p < 0.004 \)), combined with a FWE correction for multiple comparisons in small volumes (\( p_{\text{FWE}} < 0.05 \)). Because the sample studied by Hagan and colleagues included only participants younger than 18 years, we investigated group differences between patients with depression and healthy controls in the whole group, but calculated a separate analysis in participants up to age 18 years (\( n = 79 \)). For hypotheses 3 and 4, we again included age, sex and medication as covariates of no interest in the analyses.
Subgroup analyses: patients with psychiatric codiagnoses excluded
To investigate whether psychiatric codiagnoses had an impact on our data, we recalculated each analysis. To exclude effects of a history of anorexia nervosa (as found in 2 patients), we designed a model with 58 patients versus 43 controls. To exclude effects of any psychiatric codiagnosis, we excluded the 18 patients with a history of psychiatric codiagnosis (including anorexia nervosa) from calculations, setting up a model with 42 patients versus 43 controls.

Results

Behavioural data
Both groups (patients and healthy controls) included more females than males, and the majority were right-handed. One patient did not complete the BDI-II questionnaire. In the remaining group of 59 patients, the mean BDI-II score indicated a moderate degree of depressive symptoms (mean ± standard deviation = 24.31 ± 10.37). Age, handedness and sex did not differ between groups. Samples were roughly matched for education level. For more details, see Table 1.

Whole-brain analyses

Effect of age: hypothesis 1
Regression analysis revealed a significant effect of age over a large array of cortical regions (Fig. 1): the older the participant, the lower the grey-matter volumes. The global maximum in this analysis was in the paracingulate and medial prefrontal cortex (MNI coordinates: x, y, z = −8, 59, 18; Z = 5.57; number of voxels = 28181), very close to the region with the maximum effect of age on grey-matter volume found by Hagan and colleagues (MNI coordinates: x, y, z = −2, 52, 2).24 We observed further age effects in the ACC and medial prefrontal cortex, insula, lateral prefrontal cortex, inferior and superior parietal regions, and precuneus. As expected, we found no significant effects of age in the visual or somatosensory cortices, the amygdala or the hippocampus. We found no significant positive relationship between larger grey-matter volume and increasing age.

Group differences: hypothesis 2
The comparison of patients with depression and healthy controls revealed a significant group difference in the dlPFC, with greater grey-matter volumes in patients than in healthy controls (MNI coordinates: x, y, z = 44, 18, 50; Z = 4.19; number of voxels = 548; Fig. 2). We found no other significant clusters for this comparison, and testing for the opposite direction (smaller grey-matter volume in patients) provided no significant results. An interaction analysis of group × age revealed no significant results (data not shown).

Analyses according to a priori hypotheses from other studies

Differential effect in the hippocampus (ROI analysis): hypothesis 3
Comparing grey-matter volumes in the left and right hippocampus ROIs, we confirmed the finding of Jaworska and colleagues.9 We found significantly lower grey-matter volumes in the posterior hippocampus in depressed youth.

Table 1: Demographic and clinical characteristics of study participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Depressed patients (n = 60)</th>
<th>Healthy controls (n = 43)</th>
<th>Group differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, no. (%)</td>
<td>48 F (80)</td>
<td>38 F (88.4)</td>
<td>NS*</td>
</tr>
<tr>
<td>BDI-II score, mean ± SD</td>
<td>24.31 ± 10.37</td>
<td>3.07 ± 3.44</td>
<td>t&lt;sub&gt;64&lt;/sub&gt; = 14.66, p &lt; 0.001</td>
</tr>
<tr>
<td>Age, yr, mean ± SD (range)</td>
<td>17.30 ± 3.44 (13–27)</td>
<td>17.62 ± 3.85 (12–27)</td>
<td>t&lt;sub&gt;101&lt;/sub&gt; = −0.44, p = 0.66</td>
</tr>
<tr>
<td>Secondary diagnosis, no.†</td>
<td>Anorexia nervosa: 2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Social phobia: 6</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Attention-deficit/hyperactivity disorder: 4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Socialized conduct disorder: 3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Specific phobia: 2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Borderline personality disorder: 1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Handedness, no. (%‡)</td>
<td>58 right-handed (96.7)</td>
<td>40 right-handed (93)</td>
<td>NS*</td>
</tr>
<tr>
<td>Intake of psychotropic medication currently or in the past &gt; 2 months, no.§</td>
<td>Antidepressant (mirtazapine: 3; fluoxetine: 2; escitalopram: 2; sertraline: 3)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Antipsychotic (quetiapine: 1)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Psychostimulant (methylphenidate: 4)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Education level, no.¶</td>
<td>Hauptschule: 10</td>
<td>Hauptschule: 5</td>
<td>NS*</td>
</tr>
<tr>
<td></td>
<td>Realschule: 23</td>
<td>Realschule: 13</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Gymnasium: 20</td>
<td>Gymnasium: 22</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Berufsschule: 5</td>
<td>Berufsschule: 3</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Missing: 2</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

NS = not significant; SD = standard deviation.
* χ² test.
†Diagnosis according to DSM-IV; n = 18.
‡Edinburgh Handedness Inventory; a independent sample t test.
§n = 13.
¶A Gymnasium is a type of school providing secondary education, comparable to English grammar schools and United States college preparatory high schools. Hauptschule and Realschule are secondary schools. Berufsschule is a professional school that students attend along with an apprenticeship.
versus healthy controls (MNI coordinates: $x, y, z = -18, -35, -2; Z = 2.54$; number of voxels = 10).

**Group differences in the thalamus (whole-brain analysis) and the ACC (ROI analysis): hypothesis 4**

We found a differential effect in the thalamus, with greater volumes in depressed participants than in healthy controls (MNI coordinates: $x, y, z = -12, -17, 20; Z = 3.37$; number of voxels = 6) at thresholds of $p < 0.001$ at the voxel level and $p_{\text{FWE,SVC}} < 0.05$, controlling for multiple comparisons. This effect was not significant when data from participants up to age 18 years were analyzed. However, as described by Hagan and colleagues in participants up to age 18 years, and using the ROI approach we found decreased grey-matter volumes in depressed patients compared with healthy controls in the ACC, midcingular portion (MNI coordinates: $x, y, z = -6, 24, 20; Z = 3.21$; number of voxels = 72). This effect was not significant for the whole group.
Subgroup analyses: patients with psychiatric codiagnoses excluded

Exclusion of the 2 patients with a history of anorexia nervosa revealed results that were not different from those of the whole sample. Exclusion of all 18 patients with any psychiatric codiagnosis revealed unchanged results for hypotheses 1 and 3, as well as for hypothesis 4 in the thalamus. For hypothesis 2, findings in the dlPFC survived correction for multiple comparisons, but only when applying a simple cluster-level correction at $p < 0.05$ (no FWE correction). The results of these analyses are reported in greater detail in Appendix 1, available at jpn.ca/170233.

Discussion

Investigating grey-matter volumes in a comparatively large sample of peri- and postpubertal adolescents and young adults with depression, we found further support for the observation of relatively increased prefrontal volumes in this population; we also replicated findings from previous studies. Across all participants in the present study, results revealed the expected age effect (hypothesis 1): the older the participants, the lower their global grey-matter volumes, particularly in the paracingulate and the prefrontal cortex. A comparison of both groups (hypothesis 2) revealed that depressed patients had greater grey-matter volumes in the dlPFC than healthy controls. The interaction of group $\times$ age (which could support assumptions of different maturation velocities) was not significant. Replicating the findings of Jaworska and colleagues9 and Hagan and colleagues,24 our results indicated lower grey-matter volumes in the hippocampus in depressed patients (hypothesis 3) and in the ACC in depressed patients under the age of 18 (hypothesis 4) compared with healthy controls.

Our observation of increased dlPFC grey-matter volumes in depressed youth compared with healthy controls aligned well with the results of previous studies and could indicate that developmental processes in depressed adolescents follow different trajectories in brain regions known to be involved in the regulation of emotions and stress. Because increased dlPFC brain volumes are not typically seen in adults (age 24 to 35 years),33,34 have tended to support cortical (in particular, prefrontal) thinning in patients with depression, but investigations in younger groups (males average age 17 to 21.5 years) have shown thicker cortices in depressed participants compared to healthy controls in the lateral frontal cortex and in frontal regions/prefrontal cortex.35 Our results point toward a similar situation in a larger sample with respect to grey-matter volume, which is a related, but not overlapping measure of grey-matter structure.

Despite these consistencies, it must be noted that our findings were in contrast to those of a smaller study ($n = 22$ per group) by Shad and colleagues,16 which found lower grey-matter volumes in the dlPFC in depressed adolescents compared with healthy controls. One explanation for this inconsistency, besides the smaller sample size, might be that depressed participants in that study were younger (mean age 15 years; range 12 to 20 years) than those in our study (mean age 17.3 years; range 13 to 27 years). Furthermore, Hagan and colleagues,24 investigating a similarly large sample, did not report group differences in grey-matter volume between depressed and healthy adolescents. A reason for their finding might be that participants were again younger (mean age 15.65 years; range 11.83 to 17.96) than those in our study but older than those in the study of Shad and colleagues.16 These observations may guide conclusions that differences in brain maturation reflected in different prefrontal grey-matter volumes may be more prominent in slightly older adolescents. Different findings might also have been due to sex differences — Hagan and colleagues24 included more males (27.8%) than we did (20%) — and the fact that participants in that study were more often medicated (approximately 33%) than participants in our study (13.3%).

The dlPFC is an important structure when it comes to the processing of risk and fear, emotion regulation, cognitive control, monitoring of performance, response inhibition and behavioural adjustment.32,35,36 It has been suggested that maturational processes in the frontal lobe provide the neurophysiological basis for the acquisition of skills and knowledge related to higher cognitive functioning and social behaviour.37 Significant increases in the development of attentional control are seen around age 15 years,38 and development of executive functions and problem-solving abilities are found from this age into early adulthood.39 Neurobiological models of depression, hand in hand with impaired frontal lobe functioning, conceptualize some of the key phenomena observed in depression. These include biased attention to and increased processing of negative stimuli, as well as rumination that is related to impaired or dysfunctional attentional and executive control.40 In particular, cognitive behavioural therapy explicitly targets these aspects. Delayed maturational processes in the frontal lobe may be a risk factor for adolescent depression. Immature self-regulatory competence in combination with increased novelty and sensation-seeking (which have been suggested to drive commonly observed risky behaviours and emotional imbalance in adolescents) may also facilitate depression.41 Furthermore, because symptoms of depression such as rumination and compulsive behaviours have been associated with increased prefrontal volumes,27,18 neuroplastic effects may stave off the normal age-related decline and keep dlPFC volumes larger. However, it remains unclear whether delayed cortical development predisposes a person to
depression or whether depression delays the brain maturation trajectories.

As well as demonstrating greater grey-matter volumes in the prefrontal regions of depressed participants, where maturational processes peak during adolescence, we also replicated previous findings of reduced grey-matter volumes in regions such as the hippocampus.9,24 In these regions, brain maturation peaks earlier in development and is already much more advanced in the age group we investigated.11 More advanced maturational processes may be the reason why results from the hippocampus in adolescents are highly similar to those in adults.2 As in adults, prolonged exposure to stress (which induces high glucocorticoid levels and appears to harm the hippocampus in terms of neurogenesis and loss of dendritic spines45–46) has been suggested as a mechanism underlying these findings. In line with this idea, 77% of depressed minors report that a stressful life preceded or triggered the onset of their first depressive episode.45 Furthermore, Rao and colleagues46 found that early-life stress was associated with both the onset of depression and smaller hippocampal volumes. Besides that, younger age at onset of major depressive disorder has been associated with smaller hippocampi,9 and a longitudinal study noted that attenuated hippocampal growth was associated with the onset of depression.47 These findings suggest that smaller hippocampal volumes may predispose for depression, possibly because of impaired mnemonic processes,42,48 impaired executive functioning and affect regulation.49 Indeed, the meta-analysis from the ENIGMA group including data from 1728 adult patients with major depression identified smaller hippocampal volume as the most robust marker of depression, driven primarily by either earlier age of onset or recurrent depression, while recurrent depression did not moderate hippocampal volumes in early-onset patients.2 The authors concluded that early onset of depression, as in our adolescent sample, may be independently associated with lower hippocampal volumes. An alternative explanation for lower hippocampal volumes could be a lack of effective episodic foraging and information being consigned to nonconscious episodic memory, with depressed patients getting stuck on compulsive, negative perspectives on the self, world and others. Such excessive top–down rumination might limit the normal volumetric development of the hippocampus. Indeed, it has been shown that hippocampus volume accounted for impaired memory function in depressed versus healthy participants.90 Supporting this, Wang and colleagues67 found that depressed patients revealed a significant decrease of regional grey-matter volume in the parahippocampal gyrus versus healthy controls and that rumination had a mediating effect on the relationship between depression and regional grey-matter volume in the parahippocampal gyrus.

In the present study, we also found reduced ACC grey-matter volume in depressed patients versus healthy controls, again in line with the findings of Jaworska and colleagues8 and Hagan and colleagues24 in adolescents. Although the maturation peak for the ACC is less clear, it seems to occur earlier than in the dIPFC.12 Smaller ACC volumes (particularly in the subgenual portion) have been seen consistently in adults,1 but not in a longitudinal study on adolescent depression.47 The ACC is associated with conflict monitoring, social decision-making and determination of the source of social information.51,52 It also plays a key role in the processing, regulation, and appraisal of emotions.53 Therefore, reduced grey-matter volume in the ACC might relate to interpersonal and emotional impairments in depressed patients.

Finally, we found tentative evidence for aberrant thalamic volume characteristics in depressed patients versus healthy controls, similar to findings reported by Hagan and colleagues.24 The thalamus undergoes a significant amount of reorganization from early childhood through adolescence to early adulthood.50 However, structural changes in the thalamus are not consistently found in adult depression,1 despite findings of functional changes.24,55

Limitations

The present study was limited because of a skewed sex distribution biased toward female participants, which limits its generalizability to male patients. Furthermore, it was limited owing to its cross-sectional design. Although even short-term psychotropic medication might have affected brain development and volume, we were able to control only for the long term (>2 months currently or in the past), because the only available data for short-term medication in the past were inconsistent. We did not systematically explore experiences of triggering stressors such as childhood maltreatment/trauma in our study despite their well-known influence on the limbic system. Normalization of images to a standard template in voxel-based morphometry may result in some deformation of the original brain structure and possible errors in detecting small-volume differences. One strength of the present study was the inclusion of relatively young participants, which reduced the likelihood that long-term medication or chronicity of the disorder had already induced marked changes in gross brain anatomy. Another strength was the comparably large sample size obtained from a single MRI scanner.

Conclusion

We found further evidence to support different developmental trajectories in brain regions relevant for top-down processing — particularly the dIPFC — in adolescents and young adults with depression. Other brain regions, such as the hippocampus, did not show signs in support of such a rationale. Insufficient top-down control has been suggested as an explanation for the increased incidence of substance abuse, risk behaviour and affective disorders during adolescence.56 However, it remains unclear whether these changes are the cause of these disorders or an effect of already ongoing disorders. In our study, we focused on depressed participants, although some might develop other affective disorders later on (for example, initial presentation with depression is common in bipolar disorder). Longitudinal
studies in a large sample of younger adolescents are needed to shed further light on these processes and better understand natural brain maturation in healthy controls and predictors of depression and other disorders later in life. Another approach could be to calculate normative ranges of grey-matter volume in adolescents and young adults, taking into account data from healthy populations in huge data samples (e.g., ENIGMA, Human Connectome Project). In a next-step deviation, scores of depressed participants versus samples (e.g., ENIGMA, Human Connectome Project). In a next-step deviation, scores of depressed participants versus samples (e.g., ENIGMA, Human Connectome Project). In a next-step deviation, scores of depressed participants versus samples (e.g., ENIGMA, Human Connectome Project). In a next-step deviation, scores of depressed participants versus samples (e.g., ENIGMA, Human Connectome Project).

whether the brain age of depressed participants is lower than their biological age. Besides analyses of grey-matter volume, it would be interesting to focus on other aspects of grey-matter structure in cortical thickness and surface analyses.

Affiliations: From the Department of Child and Adolescent Psychiatry and Psychotherapy, Ulm University Hospital, Ulm, Germany (Straub, Brown, Bonenberger, Plener); the Department of Psychiatry and Psychotherapy III, Ulm University Hospital, Ulm, Germany (Maleko, Grön, Abler); and the Department of Child and Adolescent Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria (Plener).

Competing interests: P. Plener declares grants from Servier and Lundbeck for clinical studies outside the submitted work. No other competing interests declared.

Contributors: J. Straub, P. Plener and B. Abler designed the study. J. Straub, Brown, K. Maleko and M. Bonenberger acquired the data, which J. Straub, G. Grön and B. Abler analyzed. J. Straub, G. Grön and B. Abler wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

References


An artificial neural network model for clinical score prediction in Alzheimer disease using structural neuroimaging measures

Nikhil Bhagwat, MSc; Jon Pipitone, MSc; Aristotle N. Voineskos, MD, PhD; M. Mallar Chakravarty, PhD; Alzheimer’s Disease Neuroimaging Initiative

Background: The development of diagnostic and prognostic tools for Alzheimer disease is complicated by substantial clinical heterogeneity in prodromal stages. Many neuroimaging studies have focused on case-control classification and predicting conversion from mild cognitive impairment to Alzheimer disease, but predicting scores from clinical assessments (such as the Alzheimer’s Disease Assessment Scale or the Mini Mental State Examination) using MRI data has received less attention. Predicting clinical scores can be crucial in providing a nuanced prognosis and inferring symptomatic severity. Methods: We predicted clinical scores at the individual level using a novel anatomically partitioned artificial neural network (APANN) model. The model combined input from 2 structural MRI measures relevant to the neurodegenerative patterns observed in Alzheimer disease: hippocampal segmentations and cortical thickness. We evaluated the performance of the APANN model with 10 rounds of 10-fold cross-validation in 3 experiments, using cohorts from the Alzheimer’s Disease Neuroimaging Initiative (ADNI): ADNI1, ADNI2 and ADNI1 + 2. Results: Pearson correlation and root mean square error between the actual and predicted scores on the Alzheimer’s Disease Assessment Scale (ADNI1: $r = 0.60$; ADNI2: $r = 0.68$; ADNI1 + 2: $r = 0.63$) and Mini Mental State Examination (ADNI1: $r = 0.52$; ADNI2: $r = 0.55$; ADNI1 + 2: $r = 0.55$) showed that APANN can accurately infer clinical severity from MRI data. Limitations: To rigorously validate the model, we focused primarily on large cross-sectional baseline data sets with only proof-of-concept longitudinal results. Conclusion: The APANN provides a highly robust and scalable framework for predicting clinical severity at the individual level using high-dimensional, multimodal neuroimaging data.
Previously, computational approaches using neuroimaging measures in the context of Alzheimer disease have focused on predicting diagnosis in cross-sectional data sets, or conversion from MCI to Alzheimer disease in longitudinal analyses. However, clinicians are more likely to treat symptoms based on the results of structured assessments rather than on a specific diagnosis. In this work, we focused on predicting clinical scores of disease severity (i.e., Alzheimer’s Disease Assessment Scale [ADAS-13], Mini Mental State Examination [MMSE]) directly from neuroimaging data. Such neuroanatomically informed prediction of clinical performance at baseline and at future time points — particularly in people with MCI or significant memory concern — can help clinicians manage the clinical heterogeneity and make accurate diagnostic and prognostic decisions. Although our ultimate clinical goal is to provide longitudinal prognosis, in this report we focused primarily on a thorough validation of data sets from a single time point (baseline), an important first step in model development for longitudinal tasks. We also performed a proof-of-concept analysis to verify the ability of the proposed model to provide longitudinal prediction.

For this prediction task, we proposed an anatomically partitioned artificial neural network (APANN) model. Artificial neural networks (ANNs) and related deep-learning approaches have delivered state-of-the-art performance in classification and prediction problems for computer vision, speech recognition, natural language processing and other domains. The ANNs provide highly flexible computational frameworks that can be used to extract latent features corresponding to the hierarchical structural and functional organization of the brain and are well suited for problems with high dimensional data, unlike more standard models. To this end, the primary objective of this study was to assess whether ANN models could accurately predict ADAS-13 and MMSE clinical scores using T-weighted brain MRI data. In a larger context, we aim to build an ANN-based computational framework that can process high dimensional, distributed structural changes captured by multiple phenotypic measures to make prognostic predictions.

We designed, trained and tested our model using participants from 2 Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohorts. We used a combination of high dimensional (> 30 000) features derived from 2 neuroanatomical measures in the T-weighted images: hippocampal segmentation and cortical thickness. We generated these measures using MAGeT Brain and CIVET pipelines (see Methods), respectively. We present a model with an innovative modular design that enables the analysis of this high dimensional, multimodal input. It also allows for inclusion of new input modalities without having to retrain the entire model, and it offers simultaneous prediction of multiple clinical scores (e.g., ADAS-13 and MMSE). Given the high dimensionality of the input data, we have addressed the need for large training examples by introducing a novel data augmentation method. The method presented in this paper is not limited solely to the prediction of severity in Alzheimer disease; it can be applied to train a variety of deep-learning models that use high dimensional neuroimaging data to tackle many diagnostic and prognostic questions.

**Methods**

**Data sets**

We used baseline data from participants in the ADNI1 (n = 818) and ADNI2 (n = 788) databases (http://adni.loni.usc.edu). After exclusions based on quality control of the image preprocessing outputs, the final number of participants we used was 669 from ADNI1 and 690 from ADNI2 (see Table 1 for demographic details).

Our objective was to predict MMSE and ADAS-13 scores. The MMSE is one of the most widely used cognitive assessments for the diagnosis of Alzheimer disease and related dementias; its scores range from 0 to 30, with lower scores indicating greater cognitive impairment. The ADAS-13 is a modified version of the ADAS-cog assessment, and it has a maximum score of 85. Although ADAS-13 has some overlap with the MMSE, it also includes components that target memory, language and praxis. In contrast to the MMSE,

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ADNI1 (n = 669)</th>
<th>ADNI2 (n = 690)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acquisition</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scanner: 1.5 T</td>
<td>Scanner: 3.0 T</td>
<td></td>
</tr>
<tr>
<td>Voxel sizes: 1.2 mm × 1.25 mm × 1.25 mm</td>
<td>Voxel sizes: 1.2 mm × 1 mm × 1 mm</td>
<td></td>
</tr>
<tr>
<td><strong>Diagnosis, no.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitively healthy: 198</td>
<td>Cognitively healthy: 179</td>
<td></td>
</tr>
<tr>
<td>Late mild cognitive impairment: 326</td>
<td>Significant memory concern: 77</td>
<td></td>
</tr>
<tr>
<td>Alzheimer disease: 145</td>
<td>Early mild cognitive impairment: 162</td>
<td></td>
</tr>
<tr>
<td>Late mild cognitive impairment: 149</td>
<td>Alzheimer disease: 123</td>
<td></td>
</tr>
<tr>
<td><strong>Sex, no.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male: 377</td>
<td>Male: 361</td>
<td></td>
</tr>
<tr>
<td>Female: 292</td>
<td>Female: 329</td>
<td></td>
</tr>
<tr>
<td><strong>Age, yr</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75.0 ± 6.7</td>
<td>72.6 ± 7.2</td>
<td></td>
</tr>
<tr>
<td><strong>Education, yr</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.5 ± 3.1</td>
<td>16.3 ± 2.6</td>
<td></td>
</tr>
<tr>
<td><strong>ADAS-13 score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.4 ± 9.2 (1.0, 54.7)</td>
<td>16.1 ± 10.14 (1.0, 52.0)</td>
<td></td>
</tr>
<tr>
<td><strong>MMSE score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.7 ± 2.7 (18.0, 30.0)</td>
<td>27.5 ± 2.7 (19.0, 30.0)</td>
<td></td>
</tr>
</tbody>
</table>

ADAS-13 = Alzheimer’s Disease Assessment Scale; MMSE = Mini Mental State Examination; SD = standard deviation.

*Findings are presented as mean ± SD (minimum, maximum) unless otherwise specified.
higher scores on the ADAS-13 indicate greater cognitive impairment.

We pooled participants from all diagnostic categories to build models for the entire spectrum of clinical performance. We did not use diagnostic grouping, because we modelled Alzheimer disease progression on a continuum, a method that has been shown to be useful in other studies of Alzheimer disease progress.49,50

MRI processing

First, we preprocessed the MRIs using the bpipe pipeline (https://github.com/CobraLab/minc-bpipe-library/), consisting of N4-correction,51 neck cropping to improve linear registration and BEaST brain extraction.52 We then used the preprocessed data to extract hippocampal segmentations and cortical thickness measures, referred to as input modalities in this work. We performed computations using the GPC supercomputer at the SciNet HPC Consortium.53

Hippocampal segmentation

We produced hippocampal segmentations of $T_1$-weighted MRIs using the MAGeT brain pipeline.24,54 Briefly, this pipeline began with 5 manually segmented, high-resolution 3 T $T_1$-weighted images,55 each registered nonlinearly to 15 ADNI images selected at random (known as the template library). Then, each image in the template library was registered in a nonlinear fashion to all images in the ADNI data sets, and the segmentations from each atlas were warped via the template library transformations to each ADNI image. This process resulted in 75 (no. atlas $\times$ no. templates) candidate segmentations for each image, which were fused into a single segmentation using voxel-wise majority voting.

Cortical thickness measures

We input the preprocessed images into the CIVET pipeline29,56–59 to estimate cortical thickness at 40,962 vertices per hemisphere, which could then be grouped by region of interest (ROI) based on a surface atlas.

Anatomically partitioned artificial neural network

Artificial neural networks are a biologically inspired family of graphical machine-learning models that can perform prediction tasks using high dimensional input (Fig. 1A). These ANN models can be designed to contain multiple hidden layers, which hierarchically encode latent features that inform the objective task. The neuron connections represent a set of weights for the preceding input values, which are then combined and filtered with a nonlinear function. In neuro-imaging, a few variants of ANN models (such as autoencoders and restricted Boltzmann machines) have been investigated for classification and prediction tasks.43,60 The model used in the current study differs significantly from these approaches in both design and implementation.

From a design perspective, we leveraged the hierarchical structure of ANNs to build a modular (Fig. 1B) architecture that was capable of multimodal input integration (Fig. 1C) and multitask predictions (Fig. 1D). We achieved the following objectives in 3 stages (Fig. 1E). Stage I consisted of anatomically partitioned modules (2 hidden layers per module) that extracted features from individual anatomic input sources (hippocampus and cortical surface). These individual anatomic features served as input to stage II, where they were combined at a higher layer in the hidden-layer hierarchy. Finally, we used these integrated features to perform multiple tasks simultaneously; these task-specific hidden layers were represented by the higher layers in stage III (4 hidden layers total). This APANN mitigated overfitting by reducing the number of model parameters compared with classical, fully connected hidden-layer architectures. It also allowed for independent pretraining of each input source in a single branch. These individual pretrained branches could then be used to train stage II to integrate features efficiently.

Empirical distributions

The input dimensionality of MRI data greatly exceeds the available number of samples, leaving machine-learning models susceptible to overfitting.11,30 This necessitates the critical step of feature engineering: the transformation of high dimensional raw input to a meaningful and computationally manageable feature space.51 Techniques for addressing high dimensionality include downsampling, handcrafting features based on biological priors (e.g., atlases), principal component analysis and others. One can also increase the sample size by adding transformed data (e.g., linear transformations, image patches) to deal with the high dimensionality. In this study, we used a novel data augmentation method that leveraged the MRI preprocessing pipelines to produce a set of empirical samples for both the hippocampal and cortical thickness input modalities in place of a single point estimate per participant. This boost in training sample size made it feasible to train these models with a large parameter space and helped prevent overfitting by exposing the model to a large set of possible variations in anatomic input associated with a given severity level. Adding linear and nonlinear transformations of original input data is a common practice in machine learning.52,61 In computer vision applications, this typically means translation, rotation or dropping of certain pixels to capture a larger set of commonly encountered variations in input features to which the classifier should be invariant. In structural MRI data, we were more interested in modelling the joint voxel distribution of anatomic segmentations than in achieving high translational invariance, because the location of anatomic structures is relatively consistent across individuals. Thus, the empirical samples that were generated as part of the common segmentation and cortical surface extraction pipelines helped train the model to be invariant to the methodologically driven perturbations of input values. In turn, this mitigated overfitting and helped the model learn anatomic patterns relevant to clinical performance.

For the hippocampal inputs, the empirical samples referred to a set of “candidate segmentations” generated
An artificial neural network model for clinical score prediction

from a multi-atlas segmentation pipeline (Fig. 2A) that model the underlying joint label distribution over the set of voxels for a given participant. For the cortical thickness inputs, the empirical samples referred to cortical thickness values from a set of vertices belonging to a given cortical ROI (Fig. 2B). In traditional approaches, these samples are usually fused to produce a point estimate of the feature. We have detailed the sample-generation process for both input types below.

Hippocampal segmentation
We produced 75 candidate segmentations and 1 fused segmentation for each participant via the MAGeT brain pipeline. We segmented the ADNI1 and ADNI2 data sets using 2 separate template libraries of 15 images for each cohort. These candidate segmentations were binary masks of the left and right hippocampal voxels.

We rigidly aligned candidate segmentations to a common space (a participant chosen at random from the ADNI1 data set) to maximize anatomic correspondence across participants. We split each segmentation into left and right hemispheres and aligned both rigidly to this common space using the ANTS registration toolkit.

To remove outlier segmentations resulting from misregistration or poor segmentation, we computed the Dice \( \kappa \) between rigidly aligned candidate segmentations and the

---

![Fig. 1: (A) Structure of a generic ANN model. A neural net may consist of multiple hidden layers that encode a hierarchical set of features from input, informative of the prediction/classification task at hand. The connections between layers represent the model weights, which are updated via backpropagation based on loss function associated with the task. (B) A single feature module consisting of multiple hidden layers. This is a building block of the APANN architecture, which facilitates pretraining of individual branches per input modality. (C) A multimodal ANN with a single output task. This design consists of stage I and stage II feature modules. Stage I modules learn features from each modality that are combined in the stage II feature module. Only single-task performance is used to update the weights of the model in this architecture. (D) A multi-task ANN with a single input modality. This design consists of stage I and stage III feature modules. The stage I module learns individual features from a given modality, which are then fed into task-specific feature modules connected to the output nodes for joint prediction of the 2 tasks (ADAS-13 and MMSE score prediction). Prediction performance from both tasks is used to update the weights of the stage I feature module. Left hippocampal, right hippocampal and cortical thickness input modalities are trained separately using this design to learn input feature modules from each modality. (E) The proposed multimodal, multitask APANN model comprising anatomic partitioning. This design consists of stage I, stage II and stage III feature modules. Stage I consists of pretrained feature modules from each modality. These input features are fed into stage II to learn integrated features, which in turn are fed into the task-specific feature modules in stage III. The stage III modules are connected to the output nodes for joint prediction of the 2 tasks (ADAS-13 and MMSE score prediction). Prediction performance from both tasks is used to update the weights of the stage I and stage II feature modules. The partitioned architecture reduces the number of model parameters, which along with the pretrained feature modules helps mitigate overfitting issues. Input data dimensionality is as follows: 1686 (left hippocampal), 16471 (right hippocampal) and 686 (cortical thickness). For details regarding hyperparameters (number of hidden nodes, learning policies, weight regularization etc.) of APANN, see Table 2. ADAS-13 = Alzheimer’s Disease Assessment Scale; ANN = artificial neural network; APANN = anatomically partitioned artificial neural network; MMSE = Mini Mental State Examination; ROI = region of interest.](attachment:fig1.png)
preselected common space segmentation, and then excluded any candidate segmentations with a Dice $\kappa$ of less than 1 standard deviation from the mean over all participants.

To further compact the bounding box of all candidate segmentations, we excluded voxels with low information density by keeping only structural voxels present in at least 25% of candidate segmentations across the ADNI1 and ADNI2 data sets. After filtering operations, the 3-dimensional volumes were flattened into a 1-dimensional vector of included voxels per candidate segmentation.

Upon completion of this process, the vectorized voxels represented the hippocampal input for the APANN model. The lengths of the input vectors were 16086 for the left hippocampus and 16471 for the right.

Fig. 2: (A) Schematic of a multi-atlas segmentation pipeline depicting registration and label-fusion stages. The box highlights the candidate labels derived from different atlases that were treated as empirical samples in the context of structural labels. These labels are usually fused into a single label that serves as a point-estimate mask of a given structure. (B) Schematic of a cortical thickness estimation pipeline comprising surface registration, parcellation and average thickness estimation. The box highlights the individual vertices in a given region of interest, which are treated as empirical samples in the context of the cortical thickness measure. The thickness values of these vertices are usually averaged out to estimate mean thickness over a region of interest. CT = cortical thickness; ROI = region of interest.
Cortical thickness
Preprocessing with CIVET produces cortical thickness values at 40,962 vertices per hemisphere. We assigned these cortical vertices to unique ROIs based on a predefined atlas. We created a custom atlas (Fig. 3) consisting of 686 ROIs, maintaining bilateral symmetry (343 ROIs per hemisphere) using data-driven parcellation based on spectral clustering (http://scikit-learn.org/stable/modules/generated/sklearn.cluster.spectral_clustering.html). Spectral clustering allows for the creation of ROIs with a similar number of vertices, which is desirable for unbiased sampling of vertices to estimate cortical thickness. Also, work by Khundrakpam and colleagues suggests that increasing the spatial resolution of a cortical parcellation may improve predictive performance, further supporting the use of this data-driven atlas over neuroanatomically derived parcellations. During implementation, we used the connectivity information from the cortical mesh of the template as the adjacency matrix. Upon generating sets of vertices per ROI, we treated each vertex as a sample from a distribution that characterized the thickness of that ROI. Thus, the cortical thickness features for each individual could be characterized by a distribution of thickness values per ROI, instead of the mean thickness values computed as point estimates (Fig. 2B).

Standardization across modalities
The independent empirical sampling processes for hippocampal and cortical thickness inputs necessitated a standardization step, which is described in Appendix 1, available at jpn.ca/180016.

Training procedure
The training procedure consisted of 2 parts: training individual branches per input modality and fine-tuning the unified model consisting of pretrained branches and additional integrated and task-specific feature layers. In the first part, we trained separate models independently using individual hippocampal and cortical thickness modalities (Fig. 1D). We trained the model to jointly predict both tasks (ADAS-13 scores and MMSE scores). At the end of this training procedure, we obtained the set of weights for the hidden layers in stage I for each input branch. We then extended the model with stage II and III hidden layers and further trained it to learn integrated and task-specific feature layers (Fig. 1E). We used both tasks in this training procedure as well. For both parts, we determined the hyperparameters of the model (Table 2) using an inner cross-validation loop. The code using Caffe toolbox (http://caffe.berkeleyvision.org/) for the APANN design and training is available at https://github.com/CobraLab/NI-ML/tree/master/projects/APANN. The computational resource requirements are provided in Appendix 1.

Performance validation
We compared the performance of the APANN model separately for prediction of MMSE and ADAS-13 scores. We conducted 3 experiments to compare the performance of each cohort separately and together: ADNI1, ADNI2 and ADNI1 + 2. The latter was an effort to evaluate model robustness in a context of multicohort, multisite studies, which is becoming increasingly prevalent in the field. In each experiment, we compared the performance of the 2 inputs separately and together: hippocampal input, cortical thickness input and a combined hippocampal + cortical thickness input. We used Pearson correlation ($r$) and root mean square error (RMSE) values between true and predicted clinical scores as our
performance metrics. We evaluated all experiments using 10 rounds of a 10-fold nested cross-validation procedure. The outer folds were created by dividing the participant pool into 10 nonoverlapping subsets. During each run, we chose 9 of 10 subsets as a training set and evaluated performance on the subset that was held back. During model training, we created 3 inner folds by further dividing the training set under consideration to determine the optimal combination of hyperparameters (e.g., number of hidden nodes) using a grid search. Then, we stratified the outer folds to maintain a similar ratio of ADNI1 and ADNI2 participants in each fold. We compared the performance of APANN in all experiments against 3 commonly used machine-learning models: linear regression with lasso, support vector machine and random forest. The results are provided in Appendix 1.

Our secondary, proof-of-concept analysis consisted of a longitudinal experiment to predict clinical scores at baseline and 1 year simultaneously, using only baseline MRI data. This was in an effort to demonstrate the applicability of APANN from a clinical standpoint, where the end goal was to predict a person’s future diagnostic and/or prognostic states. We limited our analysis to the ADAS-13 scale (because its larger score range offered better sensitivity to longitudinal changes) and to the individual ADNI1 and ADNI2 cohorts. Because of missing data, the number of participants for this experiment dropped to 553 for ADNI1 and 590 for ADNI2.

Results

The mean correlation ($r$) and RMSE performance values for all 3 experiments with 3 input modality configurations are summarized in Figure 4, Table 3 and Table 4. Scatter plots for predicted and actual ADAS-13 and MMSE scores are shown in Figure 5. We generated scatter plots using scores from all test subsets in a randomly chosen round of a 10-fold run.

Results for the longitudinal experiment are shown in Figure 6. Individual results for each experiment are detailed below. Comparisons with other models are provided in Appendix 1. Briefly, results from all 3 experiments indicated that the APANN model offered better predictive performance with hippocampal inputs. The cortical thickness input, when used independently, did not offer improvement. However, the combined hippocampal + cortical thickness input offered significantly higher performance improvement over reference models across all 3 experiments.

**Experiment 1: ADNI1 cohort**

The combined hippocampal + cortical thickness input provided the best results for ADAS-13 prediction ($r = 0.60$, RMSE = 7.11). We observed similar trends for MMSE prediction with the combined hippocampal + cortical thickness input ($r = 0.52$, RMSE = 2.25). The hippocampal input alone yielded findings of $r = 0.53$, RMSE = 7.56 for ADAS-13 score prediction and $r = 0.40$, RMSE = 2.41 for MMSE. The cortical thickness input alone yielded findings of $r = 0.51$, RMSE = 7.67 for ADAS-13 score prediction and $r = 0.50$, RMSE = 2.29 for MMSE.

**Experiment 2: ADNI2 cohort**

Similar to experiment 1, the combined hippocampal + cortical thickness input provided the best results for ADAS-13 prediction ($r = 0.68$, RMSE = 7.17). We observed similar trends for MMSE prediction with the combined hippocampal + cortical thickness input ($r = 0.55$, RMSE = 2.25). The hippocampal input alone yielded findings of $r = 0.52$, RMSE = 8.32 for ADAS-13 score prediction and $r = 0.40$, RMSE = 2.51 for MMSE. The cortical thickness input alone yielded findings of $r = 0.63$, RMSE = 7.58 for ADAS-13 score prediction and $r = 0.52$, RMSE = 2.31 for MMSE.

<table>
<thead>
<tr>
<th>Model</th>
<th>Hyperparameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression with lasso</td>
<td>L1-penalty: 0.001 to 1 (with increments of 0.01)</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>Kernel: (linear, rbf), C: [0.001, 0.01, 1, 10, 100]</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>N_estimators: 10 to 210 (with increments of 25), min_sample_split: [2, 4, 6, 8]</td>
</tr>
<tr>
<td>APANN</td>
<td>Fixed hyperparameters</td>
</tr>
<tr>
<td></td>
<td>Network architecture</td>
</tr>
<tr>
<td></td>
<td>Stage I (input features): 2 hidden layers with equal nodes in each layer</td>
</tr>
<tr>
<td></td>
<td>Stage II (integrated features): 1 hidden layer</td>
</tr>
<tr>
<td></td>
<td>Stage III (task features): 1 hidden layer</td>
</tr>
<tr>
<td></td>
<td>Activation nonlinearity: ReLU</td>
</tr>
<tr>
<td></td>
<td>Tunable hyperparameters</td>
</tr>
<tr>
<td></td>
<td>Stage I number of hidden nodes: [25, 50, 100, 200]</td>
</tr>
<tr>
<td></td>
<td>Stage II number of hidden nodes: [25, 50]</td>
</tr>
<tr>
<td></td>
<td>Stage III number of hidden nodes: [25, 50]</td>
</tr>
<tr>
<td></td>
<td>Learning rate: [1e-6, 1e-5, 1e-4]</td>
</tr>
<tr>
<td></td>
<td>Learning policy: [Nesterov, Adagrad]</td>
</tr>
<tr>
<td></td>
<td>Weight decay: [1e-4, 1e-3, 1e-2]</td>
</tr>
<tr>
<td></td>
<td>Dropout rate: [0, 0.25, 0.5] (only for stage I)</td>
</tr>
</tbody>
</table>

APANN = anatomically partitioned artificial neural network; ReLU = rectified linear unit.

*We performed a grid search of the hyperparameters using a nested inner loop for each cross-validation round. For the APANN model, the fixed hyperparameters refer to a broader network of design choices that remained identical for all cross-validation rounds. The tunable hyperparameters for APANN were optimized for each fold.
Experiment 3: ADNI1 + 2 cohort

Similar to experiments 1 and 2, the combined hippocampal + cortical thickness input provided the best results for ADAS-13 prediction ($r = 0.63$, RMSE = 7.32). We observed similar trends for MMSE prediction with the combined hippocampal + cortical thickness input ($r = 0.55$, RMSE = 2.25). The hippocampal input alone yielded findings of $r = 0.54$, RMSE = 7.99 for ADAS-13 score prediction and $r = 0.45$, RMSE = 2.42 for MMSE. The cortical thickness input alone yielded findings of $r = 0.57$, RMSE = 7.79 for ADAS-13 score prediction and $r = 0.50$, RMSE = 2.37 for MMSE.

A further analysis of results in this experiment stratified by participant-cohort membership (ADNI1 v. ADNI2) showed that APANN had a smaller performance bias toward any particular cohort (i.e., models performing well on only a single cohort) than other models (see Appendix 1).

Longitudinal prediction

Similar to experiments 1 to 3, the combined hippocampal + cortical thickness input provided the best results (ADNI1: $r = 0.58$, RMSE = 7.1 for baseline and $r = 0.59$, RMSE = 9.08 for 1-year score prediction; ADNI2: $r = 0.64$, RMSE = 7.07 for baseline and $r = 0.65$, RMSE = 9.07 for 1-year score prediction). The hippocampal input alone yielded better performance than the cortical thickness input alone for baseline and 1-year score prediction in the ADNI1 cohort. The cortical thickness input alone yielded better performance than the hippocampal input alone for baseline and 1-year score prediction in the ADNI2 cohort.

![Fig. 4: Performance of anatomically partitioned artificial neural network subject to individual and combined input modalities. The Pearson $r$ and RMSE values were averaged over 10 rounds of 10 folds. All models were trained with a nested inner loop that searched for optimal hyperparameters. ADAS-13 = Alzheimer’s Disease Assessment Scale; ADNI = Alzheimer’s Disease Neuroimaging Initiative; MMSE = Mini Mental State Examination; RMSE = root mean square error.](image-url)
Discussion

We have presented an ANN model for the prediction of cognitive scores in Alzheimer disease using high dimensional structural MRI data. We showed that information from voxel-level hippocampal segmentations and highly granular cortical parcellations can be leveraged to infer cognitive performance and clinical severity at the level of the individual. This ability of the APANN model to predict ADAS-13 and MMSE and scores based on structural MRI features may prove to be valuable from a clinical perspective in helping to build prognostic tools. Our proof-of-concept longitudinal experiment showed that APANN could successfully predict future scores (at 1 year) from baseline MRI data. The results comparing APANN to several other models are provided in Appendix 1. These findings highlighted the performance gains offered by using high dimensional features as inputs. In the sections that follow, we discuss the performance of the APANN model in terms of clinical scale, input modalities, data sets and related literature.

Table 3: Prediction performance for ADAS-13 scores*

<table>
<thead>
<tr>
<th>Model</th>
<th>Hippocampal input</th>
<th>Cortical thickness input</th>
<th>Combined hippocampal and cortical thickness input</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>RMSE</td>
<td>r</td>
</tr>
<tr>
<td>ADNI1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.22 ± 0.11</td>
<td>8.72 ± 0.81</td>
<td>0.56 ± 0.08</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.23 ± 0.11</td>
<td>8.70 ± 0.85</td>
<td>0.52 ± 0.08</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.15 ± 0.10</td>
<td>9.27 ± 0.80</td>
<td>0.54 ± 0.08</td>
</tr>
<tr>
<td>APANN</td>
<td>0.53 ± 0.09</td>
<td>7.56 ± 0.76</td>
<td>0.51 ± 0.10</td>
</tr>
<tr>
<td>ADNI2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.14 ± 0.11</td>
<td>9.69 ± 0.70</td>
<td>0.61 ± 0.07</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.21 ± 0.10</td>
<td>9.75 ± 0.79</td>
<td>0.63 ± 0.07</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.24 ± 0.09</td>
<td>9.77 ± 0.76</td>
<td>0.58 ± 0.07</td>
</tr>
<tr>
<td>APANN</td>
<td>0.52 ± 0.07</td>
<td>8.32 ± 0.79</td>
<td>0.63 ± 0.07</td>
</tr>
<tr>
<td>ADNI1 + 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.12 ± 0.08</td>
<td>9.37 ± 0.50</td>
<td>0.58 ± 0.06</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.18 ± 0.07</td>
<td>9.39 ± 0.54</td>
<td>0.59 ± 0.05</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.18 ± 0.09</td>
<td>9.63 ± 0.61</td>
<td>0.57 ± 0.05</td>
</tr>
<tr>
<td>APANN</td>
<td>0.54 ± 0.06</td>
<td>7.99 ± 0.59</td>
<td>0.57 ± 0.05</td>
</tr>
</tbody>
</table>

ADNI = Alzheimer’s Disease Neuroimaging Initiative; APANN = anatomically partitioned artificial neural network; RMSE = root mean squared error; SD = standard deviation.
*Findings are presented as mean ± SD.

Table 4: Prediction performance for MMSE scores*

<table>
<thead>
<tr>
<th>Model</th>
<th>Hippocampal input</th>
<th>Cortical thickness input</th>
<th>Combined hippocampal and cortical thickness input</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>RMSE</td>
<td>r</td>
</tr>
<tr>
<td>ADNI1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.23 ± 0.12</td>
<td>2.54 ± 0.18</td>
<td>0.49 ± 0.08</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.25 ± 0.12</td>
<td>2.59 ± 0.19</td>
<td>0.48 ± 0.07</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.22 ± 0.11</td>
<td>2.63 ± 0.21</td>
<td>0.48 ± 0.08</td>
</tr>
<tr>
<td>APANN</td>
<td>0.40 ± 0.09</td>
<td>2.41 ± 0.15</td>
<td>0.50 ± 0.09</td>
</tr>
<tr>
<td>ADNI2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.19 ± 0.12</td>
<td>2.64 ± 0.19</td>
<td>0.46 ± 0.08</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.28 ± 0.14</td>
<td>2.72 ± 0.24</td>
<td>0.52 ± 0.07</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.25 ± 0.12</td>
<td>2.67 ± 0.24</td>
<td>0.50 ± 0.09</td>
</tr>
<tr>
<td>APANN</td>
<td>0.40 ± 0.09</td>
<td>2.51 ± 0.21</td>
<td>0.52 ± 0.12</td>
</tr>
<tr>
<td>ADNI1 + 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear regression with lasso</td>
<td>0.15 ± 0.08</td>
<td>2.64 ± 0.12</td>
<td>0.50 ± 0.07</td>
</tr>
<tr>
<td>Support vector regression</td>
<td>0.22 ± 0.07</td>
<td>2.71 ± 0.13</td>
<td>0.52 ± 0.07</td>
</tr>
<tr>
<td>Random forest regression</td>
<td>0.17 ± 0.08</td>
<td>2.74 ± 0.14</td>
<td>0.50 ± 0.07</td>
</tr>
<tr>
<td>APANN</td>
<td>0.45 ± 0.06</td>
<td>2.42 ± 0.14</td>
<td>0.50 ± 0.07</td>
</tr>
</tbody>
</table>

ADNI = Alzheimer’s Disease Neuroimaging Initiative; APANN = anatomically partitioned artificial neural network; MMSE = Mini Mental State Examination; RMSE = root mean squared error; SD = standard deviation.
*Findings are presented as mean ± SD.
Clinical scale comparisons

Performance comparisons between clinical scales based on correlation values indicated that predicting MMSE scores was more challenging across all inputs and cohorts. This disparity between performances may have been due to the higher sensitivity of the ADAS-13 assessment, reflected in its comparatively larger scoring range, which improved its association with the structural measures.

Input modality comparisons

The results from all 3 experiments indicated that the APANN model offered better predictive performance with the combined hippocampal + cortical thickness input. The use of cortical thickness outperformed hippocampal segmentation in all 3 experiments for both scales, except in the ADNI1 cohort for ADAS-13 prediction, where the hippocampal segmentation input showed a slightly higher performance. This finding highlighted the importance of incorporating multiple phenotypes for biomarker development that are indicative of cognitive performance. The ability of the APANN model to handle multimodal input is crucial for building clinical tools to leverage disparate MRI, clinical and genetic markers.

Data set comparisons

Between experiments 1 and 2, we observed that the ADNI2 cohort yielded better performance than the ADNI1 cohort across all models. This may have been because of the differences in acquisition protocols, because ADNI2 images were acquired at a higher field strength with better resolution. Such an improvement in image acquisition would likely provide superior quality segmentations and cortical thickness measures. In experiment 3, we combined data from the ADNI1 and ADNI2 cohorts. Pooling data from different data

---

**Fig. 5:** Scatter plots for predicted and actual ADAS-13 and MMSE scores for 3 cohorts (ADNI1, ADNI2 and ADNI1 + 2). Scatter plots were generated by concatenating scores from all the test subsets of a randomly chosen 10-fold validation run. ADAS-13 = Alzheimer’s Disease Assessment Scale; ADNI = Alzheimer’s Disease Neuroimaging Initiative; MMSE = Mini Mental State Examination.
Fig. 6: Simultaneous predictions of ADAS-13 scores at baseline and 1 year. The top 2 rows show the Pearson $r$ values based on predicted and actual ADAS-13 scores over 10-fold cross-validation for ADNI1 and ADNI2 cohorts, respectively. The bottom 2 rows show the RMSE between predicted and actual ADAS-13 scores for ADNI1 and ADNI2 cohorts, respectively. The left column shows performance at baseline, and the right column shows performance at 1 year. Models were trained separately for each input. ADAS-13 = Alzheimer’s Disease Assessment Scale; ADNI = Alzheimer’s Disease Neuroimaging Initiative; APANN = anatomically partitioned artificial neural network; RMSE = root mean square error.
sets is increasingly important to verify the generalizability of the model in a larger population that extends beyond a single study. Interestingly, experiment 3 outperformed experiment 1, but underperformed compared with experiment 2. This was partially expected because of substantial differences in the individual feature distributions (e.g., hippocampal segmentations) resulting from differences in the acquisition protocols. In such cases, it becomes imperative to build models that are invariant to data set-specific biases resulting from nonuniform data-collection practices. The results from experiment 3 showed that APANN offered consistent performance that was comparable to that of experiments 1 and 2, and it had low data set-specific bias compared with other models (see Appendix 1). We speculate that models incorporating high dimensional, multimodal input were less susceptible to multicohort and multisite study-design artifacts, a characteristic that is desirable for the development of clinical tools in practical settings.

**Longitudinal analysis**

Consistent with the first 3 experiments, the combined hippocampal segmentation + cortical thickness input offered the best performance for 1-year score prediction, with similar correlation results but higher RMSE. This finding suggests that uncertainty is likely to increase with a larger time span for longitudinal tasks (1 year v. 2 years v. 5 years), making predictions more challenging. As well, further consideration is needed of cases in which information from multiple time points (baseline + 1 year) is used to generate subsequent (2 years +) performance prediction. Missing time points become an increasingly important barrier to such tasks. Nevertheless, APANN showed promising results for investigating more sophisticated longitudinal predictions.

**Related work**

Prediction of clinical scores is a relatively underexplored task. For a fair comparison, we have limited our discussion to 2 recent studies involving baseline prediction with MRI features by Stonington and colleagues and Zhang and colleagues. Both works used structural MRI from the ADNI1 baseline data set to predict MMSE and ADAS-cog scores (which uses 11 of the 13 subscales of ADAS-13; http://adni.loni.usc.edu/data-samples/data-faq/). The ADAS-cog and ADAS-13 scores are strongly correlated \((r > 0.9\) for the ADNI1 and ADNI2 cohorts considered in this manuscript). Stonington and colleagues used relevance vector regression models with a sample size of 586; correlation values were \(r = 0.48\) (MMSE) and \(r = 0.57\) (ADAS-cog). Zhang and colleagues proposed a computational framework called Multi-Modal Multi-Task (M3T) that offers multitask feature selection and multimodal support vector machines for regression and classification tasks. With only MRI-based features, M3T achieved correlations of \(r = 0.50\) (MMSE) and \(r = 0.60\) (ADAS-cog) with a sample size of 186. In comparison, the APANN model offered correlations of \(r = 0.52\) (MMSE), and \(r = 0.60\) (ADAS-13) with a much larger cohort (669 ADNI1 participants). Although APANN offered similar performance for the ADNI1 data set, it had several key advantages over the other models. In contrast to M3T, which implemented 2 separate stages for feature extraction and regression (or classification) tasks, APANN provided a unified model that performed seamless feature extraction and multitask prediction using multimodal input. From a scalability perspective, APANN was capable of handling high dimensional input and extending to incorporate new modalities without retraining the entire model. In contrast, M3T had 93 magnetic resonance atlas-based features with a total of 189 multimodal (MRI, FDG-PET and cerebrospinal fluid) features. Moreover, with APANN we replicated performance in the ADNI2 cohort and demonstrated an improved correlation performance of \(r = 0.55\) (MMSE) and \(r = 0.68\) (ADAS-13) with 690 participants, further validating the model’s generalizability.

Other recent works have addressed clinical score prediction using sparse Bayesian learning and graph-guided feature selection, with 98 and 93 imaging features, respectively. Both works reported strong performance in Alzheimer disease and cognitively normal groups, but performance degraded after inclusion of people with MCI. For example, Yu and colleagues reported correlations of \(r = 0.745\) (MMSE) and \(r = 0.74\) (ADAS-cog) for specific subsets of Alzheimer disease/cognitively normal participants, but performance degraded to \(r = 0.382\) (MMSE) and \(r = 0.472\) (ADAS-cog) for a subset of MCI/cognitively normal participants. Clinically, the prognosis of people with MCI is of high interest. Predicting their cognitive performance is crucial for early interventions, potential lifestyle changes and treatment planning. To the best of our knowledge, APANN is the first work to tackle high input dimensionality (> 30,000 features), validated across the continuum from healthy controls to patients with Alzheimer disease, in multiple cohorts with site and scanner differences. Such validation is increasingly important with the availability of newer and larger data sets, such as the UK biobank (www.ukbiobank.ac.uk/about-biobank-uk/).

**Clinical translation**

The ultimate clinical goal of this work is to provide longitudinal prognosis and to predict individuals’ future clinical states. The rigorously validated APANN provides a computational platform for a variety of longitudinal tasks, such as the 1-year ADAS-13 prediction task investigated in the proof-of-concept experiment. We envision the APANN model applied to the MRI data of people at risk from prodromal stages (MCI, significant memory concern etc.) and even early stages of Alzheimer disease to predict their future clinical scores and other clinical-state proxies. The ability of the APANN model to capture relevant subtle neuroanatomical changes from high dimensional, multimodal MRI data can be leveraged to provide nuanced diagnosis and prognosis for various symptom subdomains, assisting or verifying clinicians’ decision-making. Having a clear prognosis can help with early intervention, clinical trial recruitment and caregiver arrangements.
Limitations

In this work we applied APANN primarily to cross-sectional data sets and a proof-of-concept longitudinal data set. From a clinical perspective, it is crucial to note that the use of a specific clinical or cognitive test is subjective, contingent on availability and associated with its own set of biases. Further, similar to the clinical diagnosis that uses several sources of information to create a composite of the patient’s clinical profile, we envision the proposed MRI-based prediction framework as another assistive instrument that will be interpreted in the larger context of an overall clinical picture. We acknowledge that the cross-sectional experiments in this work were a first step toward building assistive MRI-based models. We believe that the design flexibility of APANN can be used for handling multimodal input and multiple scale predictions that could minimize modality-specific and scale-specific biases, respectively.

Large-scale models such as APANN that are subjected to high dimensional input require significant computational resources. Thus, we have limited the scope of this work to classical ANNs as a prototypical example to demonstrate the feasibility of large-scale analysis with structural neuroimaging data. Nevertheless, the training regimens discussed in this work should motivate further development of state-of-the-art neural network architectures, such as 3-dimensional convolutional networks, toward various neuroimaging applications. Another common drawback of models with deep architectures is the lack of interpretability of the model parameters compared with simpler models; this prohibits localizing most predictive brain regions. In our view, this limitation is a model design trade-off that in turn allows for the capture of distributed changes that are often present in the heterogeneous atrophy patterns of Alzheimer disease proclivity. The computational flexibility of ANNs allow us to model the collective impact of these atrophy patterns and predict clinical performance more accurately.

Conclusion

The presented APANN model, together with empirical sampling procedures, offers a sophisticated machine-learning framework for high dimensional, multimodal structural neuroimaging analysis. By going beyond low-dimensional, anatomic prior-based feature sets, we can build more sensitive models capable of capturing the subtle neuroanatomical changes associated with cognitive symptoms in Alzheimer disease. The results validate the strong predictive performance of the APANN model across 2 independent cohorts, as well as its robustness when these 2 cohorts were combined. From a clinical standpoint, these attributes make APANN a promising approach for building diagnostic and prognostic tools that would help identify people at risk and provide clinical-trajectory assessments, facilitating early intervention and treatment planning.

Acknowledgements: N. Bhagwat receives support from the Alzheimer Society of Canada. A. Voinoskos is funded by the Canadian Institutes of Health Research, the Ontario Mental Health Foundation, the Brain and Behavior Research Foundation and the National Institute of Mental Health (R01MH099167 and R01MH102324). M. Chakravarty is funded by the Weston Brain Institute, the Alzheimer Society of Canada, the Michael J. Fox Foundation for Parkinson’s Research, the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and Fondation de Recherches Santé Québec. Computations were performed on the GPC supercomputer at the SciNet HPC Consortium and the Kimel Family Translational Imaging-Genetics Research (TIGR) Lab computing cluster. SciNet is funded by the Canada Foundation for Innovation under the auspices of Compute Canada, the Government of Ontario, the Ontario Research Fund Research Excellence Program and the University of Toronto. The TIGR Lab cluster is funded by the Canada Foundation for Innovation Research Hospital Fund. Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904), and ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from Abbott; the Alzheimer’s Association; the Alzheimer’s Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research Development LLC; Johnson & Johnson Pharmaceutical Research Development LLC; Medpace Inc.; Merck & Co. Inc.; Meso Scale Diagnostics LLC; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research provides funds to support ADNI clinical sites in Canada. Private-sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is Rev March 26, 2012, coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. The ADNI data are disseminated by the Laboratory for Neuroimaging at the University of California, Los Angeles. This research was also supported by National Institutes of Health grants P30 AG010129 and K01 AG030514.

Affiliations: From the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ont. (Bhagwat, Chakravarty); the Cerebral Imaging Centre, Douglas Mental Health University Institute, Verdun, Que. (Bhagwat, Chakravarty); the Kimel Family Translational Imaging-Genetics Research Lab, Research Imaging Centre, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ont. (Bhagwat, Pipitone, Voinoskos); the Department of Psychiatry, University of Toronto, Toronto, Ont. (Voinoskos); and the Department of Psychiatry, McGill University, Montreal, Que. (Chakravarty), Canada.

Competing interests: None declared.

Contributors: N. Bhagwat, J. Pipitone and M. Chakravarty designed this work. Data were collected by the Alzheimer’s Disease Neuroimaging Initiative, and all authors participated in data analysis. N. Bhagwat and J. Pipitone wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

References

An artificial neural network model for clinical score prediction

2011;56:766-81.

2016;87:235-43.

dementia using latent class analysis.

identification of cognitive subtypes in Alzheimer's disease

atrophy: longitudinal trajectories and clinical applications.

2013;8:e77949.

assessment using large-scale machine learning methods.

multiple automatically generated templates.

segmentation of the whole hippocampus and subfields using


semantic dementia.

volumetry in mild cognitive impairment, Alzheimer's disease and

2009;461:916.

Neurol-

based hippocampal volume in mild cognitive impairment.

Braak H, Braak E. Frequency of stages of Alzheimer-related lesions


comparison of ten methods using the ADNI database.

of patients with Alzheimer's disease from structural MRI: a

Mueller SG, Weiner MW. Selective effect of age, Apo e4, and

Alzheimer’s disease on hippocampal subfields. Hippocampus

2009;19:558-64.


of the fornix, fimbria, and alveus on high-resolution 3T MRI:

application via fully-automated mapping of the human memory

circuit white and grey matter in healthy and pathological aging.


thickness in Alzheimer’s disease identified by computational


using expert priors: application to hippocampus and ventricle


Sterin Y. Cognitive reserve in ageing and Alzheimer’s disease.


Whitwell JL, Dickson DW, Murray ME, et al. Neuroimaging

correlates of pathologically defined subtypes of Alzheimer’s

Lam B, Masellis M, Freedman M, et al. Clinical, imaging, and

pathological heterogeneity of the Alzheimer’s disease syndrome.


Eskildsen SF, Poupon C, Garcia-Lorenzo D, et al. Prediction of

Alzheimer’s disease in subjects with mild cognitive impairment

from the ADNI cohort using patterns of cortical thinning.


Braak H, Braak E. Frequency of stages of Alzheimer-related lesions


subtypes of Alzheimer’s disease based on patterns of brain

atrophy: longitudinal trajectories and clinical applications. Sci Rep

2017;7:46265.

discovery in neuroimaging data: finding a needle in a haystack.


Scheltens NME, Galindo-Garre F, Pijnenburg YAL, et al. The

identification of cognitive subtypes in Alzheimer’s disease
dementia using latent class analysis. J Neural Neurosurg Psychiatry

2016;87:235-43.

Jack CR Jr, Petersen RC, Xu YC, et al. Prediction of AD with MRI-

based hippocampal volume in mild cognitive impairment. Neurology


change in mild cognitive impairment to baseline MRI features.


Frisoni GB, Fox NC, Jack CR Jr, et al. The clinical use of structural


Sabuncu MR, Desikan RS, Sepulcre J, et al. The dynamics of cortical

and hippocampal atrophy in Alzheimer disease. Arch Neurol


Neuroimaging Initiative. Simultaneous segmentation and grading

of anatomical structures for patient’s classification: application to


Subscriptions and sales
Annual subscriptions are available at the following rates in 2019 (Canadian customers pay in Canadian funds and add applicable taxes; US and rest of world pay in US funds): Can$256 or US$295 individuals, Can$393 or US$459 institutions, Can$50 for Canadian students or residents, and Can$38 or US$49 for single or back issues. For more information or to subscribe, please contact the Joule Inc. Subscription Office, PO Box 830350, Birmingham AL 35283-0350; phone 800 633-4931 (Canada and Continental US only) or 205 995-1567; fax 205 995-1588; cma@subscriptionoffice.com.

Change of address
Subscribers should send the new address and the effective date of change to the CMA Member Service Centre at cmamsc@cma.ca. Allow at least 8 weeks’ notice to ensure uninterrupted service.

Reprints
Commercial reprints and e-prints, and author reprints can be purchased from Isabelle Laurendeau, business development manager, at isabelle.laurendeau@cma.ca, 800-663-7336 x 8441.

Permissions
Copyright for all material is held by Joule Inc., a wholly owned subsidiary of the Canadian Medical Association. We are a member of Access Copyright, The Canadian Copyright Licensing Agency, who may grant organizations and individuals, on our behalf, the right to respond to permissions requests. Please submit your request to Access Copyright using the online permission request service at http://discovery.accesscopyright.ca. For more information on licensing or to obtain a price estimate, visit www.accesscopyright.ca/permissions/

Microform, abstracting and indexing
The journal is abstracted or indexed in: MEDLINE/Index Medicus, Current Contents, Biossearch Index, Social Sciences Citation Index, Biological Abstracts, Embase/Excerpta Medica, e-psyche, Mental Health Abstracts, ISI Web of Science, Child Development Abstracts and Bibliography, Standard Periodical Directory, Arts and Humanities Citation Index, Research Alert (formerly Automated Subject Citation Alert), Criminal Justice Abstracts, Neuroscience Citation Index, Psychological Abstracts, Sociological Abstracts. JPN is available on microform from ProQuest Information and Learning, PO Box 134, Ann Arbor MI 48106-1346; il.proquest.com.
Changes of motor cortical excitability and response inhibition in patients with obsessive–compulsive disorder

Jee In Kang, MD, PhD; Deog Young Kim, MD, PhD; Chang-il Lee, MD; Chan-Hyung Kim, MD, PhD; Se Joo Kim, MD, PhD

Introduction

Obsessive–compulsive disorder (OCD) is a neuropsychiatric disorder characterized by distressing, time-consuming or impairing obsessions (recurrent thoughts, images or urges) and compulsions (repetitive behaviours or mental acts). Several lines of evidence support the concept that dysfunctions of inhibitory control in frontostriatal circuits are associated with the inability to inhibit cognition and behaviour, as well as increased action-monitoring in people with OCD. Behavioural studies have shown that people with OCD have impaired motor and cognitive inhibitory mechanisms. People with OCD have shown impaired response inhibition in behavioural inhibition tasks such as the go/no-go task, and young people with OCD have revealed deficits of behavioural inhibition in an oculomotor task. In addition, neuroimaging studies have shown that excessive activation in specific brain regions, including the anterior cingulate and orbitofrontal cortices, in patients with OCD during trials required response inhibition. Accumulating evidence has suggested that altered cortical inhibition, such as an imbalance of direct and indirect feedback loops within cortical–striatal–thalamic–cortical (CSTC) circuits, may contribute to the characteristic cognitive disruptions of OCD. An optogenetic mouse model has shown that stimulation in the frontostriatal pathway can alleviate OCD-relevant behaviours. These findings indicate that deficits in cortical inhibition in motor and cognitive processes may play a key role in the mechanisms of behavioural inhibition deficit and symptom formation in patients with OCD.

Paired-pulse transcranial magnetic stimulation (TMS) is a noninvasive technique that allows researchers to directly assess cortical excitability, which depends on the balance between excitatory and inhibitory circuits. Paired-pulse TMS with subthreshold conditioning can test cortical excitability at interstimulus intervals (ISIs) of 1 ms to 4 ms and intracortical facilitation (ICF) at longer ISIs of 6 ms to 20 ms.
A growing body of evidence suggests that SICI is mediated by \( \gamma \)-aminobutyric acid (GABA)-A receptors,\(^{14}\) while ICF depends, in part, on glutamatergic neurotransmission.\(^{15,16}\) Another key measure of TMS is the cortical silent period (CSP), which reflects GABA-mediated motor cortical postsynaptic inhibition. It has been suggested that short CSPs elicited by low stimulus intensity are associated with the activation of GABA-A receptors, while long CSPs elicited by high stimulus intensity are associated with the activation of GABA-B receptors.\(^{17}\) These TMS parameters can be a promising neurophysiological biomarker for elucidating the underlying mechanisms of psychiatric disorders that involve GABAergic and glutamatergic neurotransmission.

To date, inhibitory deficits and enhanced intracortical facilitation have been implicated in OCD, with a limited number of studies examining cortical excitability using paired-pulse TMS.\(^{18}\) One study reported that in a small OCD sample, patients showed significantly reduced SICI and a decreased motor threshold compared with healthy controls.\(^{19}\) In contrast, a subsequent study in a healthy population reported that decreased SICI may be more linked to anxiety and depression personality traits than to OCD itself.\(^{20}\) Another study in patients with OCD (\( n = 34 \)) showed that OCD was associated with shortened CSP and increased ICF, and not associated with SICI, suggesting that dysregulation of GABA-B and \( N \)-methyl-\( \beta \)-aspartate (NMDA) receptor-mediated neurotransmission may be involved in the pathophysiology of OCD.\(^{21}\)

We aimed to investigate differences in cortical excitability between patients with OCD and healthy controls using paired-pulse TMS. We also examined associations between TMS indices and clinical characteristics — including age of onset and response inhibition in the go/no-go paradigm — to examine whether altered cortical excitability contributes to symptom formation and behavioural inhibition deficit in patients with OCD. We hypothesized that patients with OCD would have abnormal cortical excitability with reduced resting motor threshold (RMT), shortened CSP, reduced SICI or increased ICF and would show correlations between neurophysiological alteration and specific OCD-related characteristics.

**Methods**

**Participants**

We recruited 55 patients with OCD from the OCD clinic in Severance Hospital and Yonsei Phil Neuropsychiatric Clinic, and 42 age-matched healthy controls via advertising. Participants underwent a face-to-face interview based on the Structured Clinical Interview for DSM-IV disorders.\(^{22}\) Inclusion criteria for the patient group were age 18 to 50 years, a current DSM-IV diagnosis of OCD and stable maintenance of medication for at least 8 weeks before enrollment. Patients with comorbid depression but without psychotic features were included in the study only if the obsessive–compulsive symptoms were their most prominent symptoms and the onset of OCD predated the onset of depression. Controls were included if they had been physically healthy and medication-free for the past 6 months, had no history of psychiatric disorders (including OCD and depressive disorders) and no family history of psychiatric disorders among first-degree relatives. All participants were right-handed. Participants were excluded if they presented with a movement disorder other than a tic; any psychotic symptoms; other anxiety disorders; an intellectual disability; alcohol or other substance abuse within the last 6 months; or a history of seizure, psychosurgery, encephalitis or significant head trauma. The study was approved by the ethical committee of Severance Hospital, and written informed consent was obtained from all participants.

Among the recruited participants, 2 patients and 2 controls with comorbid psychiatric problems were excluded. Two patients and 1 control who felt uncomfortable during the TMS procedure and wanted to stop participation (2 headache, 1 nonspecific discomfort) were also excluded because of incomplete acquisition of TMS data. Ultimately, 51 patients with OCD (mean age ± standard deviation = 27.43 ± 7.64 years) and 39 age-matched healthy controls (27.36 ± 6.99 years) were included in our analyses. The present sample exceeded the recommended sample size of 34 in each group required to detect the global effect with medium effect size in a multivariate analysis of variance according to G*Power.\(^{23}\)

**TMS protocol and stimulation parameters**

For TMS, we used 2 Magstim-200 stimulators connected via a Bistim module with a 70 mm figure-8 coil (Magstim Company Ltd.). We took TMS recordings with surface electrodes from the abductor digitii minimi muscle. The coil was held with the grip pointing backward and perpendicular to the central sulcus. For optimal coil positioning, we measured the amplitudes of the motor evoked potential in the resting abductor digitii minimi by moving the coil in 1 cm steps over the presumed area of the contralateral motor cortex. We performed the paired-pulse paradigms on both hemispheres.

For measures of motor cortex excitability, we examined motor evoked potential, RMT, CSP, SICI and ICF. We determined RMT over the primary motor cortex in both groups by finding the minimal intensity required to elicit at least 5 motor evoked potentials of 50 mV out of 10 stimulations of the contralateral abductor digitii minimi muscle.

We obtained CSP by applying stimuli with an intensity of 40% above the active motor threshold. We applied and recorded TMS pulses with 10 trials at 140% of the active motor threshold during a low-level voluntary contraction of the abductor digitii minimi muscle. The duration of the CSP was defined in the rectified single trials as being from the end of the preceding motor evoked responses to the return of the amplitude of the mean voluntary electromyographic activity before TMS. The durations of each CSP elicited from 10 consecutive electromyographic signals were rectified and then averaged.

We used paired-pulse TMS with subthreshold conditioning to test SICI with ISIs of 2 ms and 3 ms and ICF with ISIs of 10 ms and 15 ms. The conditioning stimulus was set at 80% of the RMT and preceded the test stimulus (110% to 120% of...
the RMT), which produced a response of about 1 mV peak-to-peak amplitude. We applied 10 paired-pulse TMS trials of each ISI in 4 randomly intermixed conditions. Time between trials was 5 s. We measured the peak-to-peak amplitudes and then averaged them.

**Measures of clinical symptoms and traits**

We assessed all participants for obsessive-compulsive symptoms using the Maudsley Obsessive Compulsive Inventory (MOCI). The MOCI is a self-rating instrument of 30 dichotomous items designed to measure obsession and compulsion symptoms. It consists of 4 subscales: checking, washing, doubting and slowness. We also assessed patients with OCD using the Yale–Brown Obsessive Compulsive Scale (Y-BOCS) for the severity of obsessive-compulsive symptoms. We assessed depressive symptoms and anxiety levels using the Montgomery–Åsberg Depression Rating Scale (MADRS) and the Hamilton Anxiety Rating Scale (HARS), the most widely used semi-structured assessment scales, administered by trained psychiatrists. Missing values below 5% were replaced by an expectation-maximization algorithm.

**Go/no-go test**

To measure motor response inhibition, we performed the computerized go/no-go test (a faster variant of the classical go/no-go paradigm) in patients with OCD. We have previously reported impaired response inhibition in patients with OCD compared with healthy controls, using the go/no-go test. The task requires selection of a response (indicated by a “go” signal) or no response (indicated by a “no-go” signal). Patients with OCD were asked to respond to go signals (airplanes) appearing on the centre of the screen but not to no-go signals (bombs). The task was administered in 2 blocks, with 90 practice trials and 180 testing trials (126 go trials and 54 no-go trials) in randomized order. Patients were asked to inhibit their motor response when the no-go signals appeared on the screen. The ISI was 1000 ms, including a stimulus duration of 200 ms followed by a blank screen for 800 ms. The dependent variables of the inhibitory process were the percentage of successful inhibition trials and the mean reaction time for the correct go trials (ms).

**Statistical analysis**

We performed a priori power analysis using G*Power 3.1. A sample size of 68 participants (minimum 34 in each group) was needed to potentially detect a global effect of multivariate analysis of variance with a medium effect size (f² = 0.25), a power of 0.9 and 4 response variables.

We analyzed data using SPSS 23.0 for Windows (SPSS Inc.). Significance was set at p < 0.05. All tests were 2-tailed. We used the t test to evaluate differences in demographic and clinical characteristics between groups for continuous variables and χ² tests for categorical variables.

We performed repeated-measures multivariate analysis of covariance (MANCOVA) to compare neurophysiological indices of TMS between patients with OCD and healthy controls, with hemisphere (left v. right) as the repeated factor, group (OCD v. controls) as the between-subjects factor, and TMS indices of RMT, CSP, SICI and ICF as the within-subjects factors. In MANCOVA, we used the mean values of the conditioned motor evoked potential size (a ratio of the conditioned motor evoked potential amplitude to the control motor evoked response) at ISIs of 2 ms and 3 ms as the parameter for SICI, and 10 ms and 15 ms for ICF. Covariates included age and sex to remove any possible effects on cortical excitability, based on previous findings. Significant results from MANCOVA were followed by separate univariate ANCOVA. We calculated η² values as a measurement of effect size, considering that a η² of 0.01 was small, 0.04 moderate and 0.1 large. For post-hoc ANCOVA, we adjusted the significance threshold using a Bonferroni approach to correct for tests of 4 dependent variables (i.e., 0.05/4 = 0.0125).

We also used partial correlations (pr) with covariates to explore the relationships between neurophysiological indices and clinical variables, such as onset age and parameter, regarding the inhibitory function on the go/no-go test (mean reaction time on the correct go trials) in patients with OCD.

**Results**

We found no significant differences between patients with OCD and controls in terms of age, sex or educational level. All patients were undergoing treatment with serotonin reuptake inhibitors (SRIs). Clinical characteristics are presented in Table 1.

Repeated-measures MANCOVA with hemisphere (left v. right) as the repeated factor, diagnostic status of OCD as a between-subjects factor, age and sex as covariates, and TMS variables (RMT, CSP, ICF and SICI) as the within-subjects factors showed a significant difference between patients with OCD and healthy controls (F1,46 = 10.66; p < 0.001; η² = 0.339). We found no significant main effect of hemisphere (left v. right; F1,46 = 1.140; p = 0.34). We also observed no significant interaction effects between age and hemisphere (F1,46 = 0.355; p = 0.70), sex and hemisphere (F1,46 = 1.542; p = 0.20) or OCD diagnosis and hemisphere (F1,46 = 0.722; p = 0.58). In post hoc between-group comparisons, patients with OCD showed significantly shortened CSP compared to healthy controls (F1 = 12.604; p = 0.001; η² = 0.128; Fig. 1). Patients with OCD also had significantly decreased ICF (F1 = 10.298; p = 0.002; η² = 0.107). We found no significant difference in SICI (F1 = 3.579; p = 0.052; η² = 0.043) or RMT (F1 = 0.073; p = 0.79) between groups (Fig. 1).

After excluding patients with comorbid depression (n = 21), MANCOVA with the OCD group (n = 30) also showed that the between-group effects of CSP (F = 5.678; p = 0.020; η² = 0.080) and ICF (F = 11.798; p = 0.001; η² = 0.154) were still meaningful.

In the partial correlation analyses with age and sex as covariates, we found positive correlations at the trend level between CSP and mean reaction time in the go/no-go test.
OCD was consistent with previous results.\(^{19,21,33}\) Furthermore, it involves the largest sample of patients with OCD among studies using the paired-pulse TMS paradigm. Our results, which compared controls, patients with OCD and healthy controls, showed that patients with OCD had altered cortical excitability in terms of shortened CSP and decreased ICF. We found no significant differences in RMT and SICI between patients with OCD and controls. These findings suggest that alterations of inhibitory neurotransmission mediated by GABA-B receptors and excitatory neurotransmission mediated by NMDA receptors may be involved in the pathophysiology of OCD.

Discussion

The present study examined cortical excitability using paired-pulse TMS in patients with OCD and healthy controls to elucidate cortical inhibitory deficits in OCD. To our knowledge, it involves the largest sample of patients with OCD among studies using the paired-pulse TMS paradigm. Our results show that compared with controls, patients with OCD had altered cortical excitability in terms of shortened CSP and decreased ICF. We found no significant differences in RMT and SICI between patients with OCD and controls. These findings suggest that alterations of inhibitory neurotransmission mediated by GABA-B receptors and excitatory neurotransmission mediated by NMDA receptors may be involved in the pathophysiology of OCD.

Our finding of shortened CSP, a major inhibitory index, in OCD was consistent with previous results.\(^ {19,21,33}\) Furthermore, in our correlation analyses, shortened CSP was associated with prompt reaction time in a go/no-go task and early onset of OCD. In particular, after adjusting for depressive symptoms, the correlation between CSP and reaction time in the go/no-go task was significant (\(p = 0.013\)). Shortened CSP in patients with OCD along with a shorter reaction time in go trials may indicate that the neurophysiological mechanism of impaired cortical inhibition plays a crucial role in impaired inhibitory control of thoughts and behaviours in OCD. In addition, because CSP at high stimulus intensity is considered a marker of GABA-B receptor-mediated inhibitory function,\(^ {34}\) these findings in patients with OCD suggest that a lack of cortical inhibition via GABA-B receptor-mediated dysregulation might contribute to response inhibition and acceleration of OCD onset. However, since the present study had a cross-sectional design, further prospective research is needed to prove a possible association for OCD onset and response inhibition.

Regarding SICI, another main inhibitory parameter of TMS, our results did not show any significant finding of reduced cortical inhibition in patients with OCD compared with controls. This finding was consistent with that of Richter and colleagues,\(^ {25}\) who found no differences between patients with OCD (\(n = 34\)) and healthy controls. However, it was inconsistent with a recent finding by Khedr and colleagues,\(^ {33}\) which showed that patients with OCD (\(n = 45\)) had significantly reduced SICI compared with healthy controls. Greenberg and colleagues\(^ {19}\) also showed that SICI in patients with OCD (\(n = 16\)) was significantly lower than in healthy controls. These discrepancies may have been partially due to the low statistical power of small sample sizes and medication effects. Since SRIs are known to modulate GABA release\(^ {36}\) and enhance SICI,\(^ {36}\) the possible effect of SRIs could have concealed any potential SICI deficits in the current study. To confirm the characteristics of cortical excitability in OCD, replication by future studies using paired-pulse TMS in larger samples with drug-naïve patients is needed.

Contrary to our expectations, for the facilitatory component of intracortical excitability, patients with OCD had

---

**Table 1: Demographic and clinical characteristics between patients with OCD and healthy controls**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Healthy controls (n = 39)</th>
<th>OCD (n = 51)</th>
<th>t or (\chi^2)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/F, n</td>
<td>30/9</td>
<td>40/11</td>
<td>0.029</td>
<td>0.87</td>
</tr>
<tr>
<td>Education, yr</td>
<td>14.10 ± 2.00</td>
<td>13.41 ± 2.19</td>
<td>–1.54</td>
<td>0.13</td>
</tr>
<tr>
<td>MOCI score</td>
<td>4.61 ± 2.67</td>
<td>17.49 ± 5.90</td>
<td>13.85</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age of OCD onset, yr</td>
<td>15.90 ± 5.80</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Y-BOCS score</td>
<td>23.51 ± 7.23</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>MADRS score</td>
<td>16.94 ± 10.02</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>HARS score</td>
<td>14.45 ± 9.95</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Comorbid depression, n</td>
<td>0</td>
<td>21</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Receiving SRIs, n</td>
<td>0</td>
<td>51†</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzodiazepines, n</td>
<td>0</td>
<td>28‡</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Antipsychotics, n</td>
<td>0</td>
<td>4§</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Go/no-go task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Successful inhibition trials, %</td>
<td>—</td>
<td>84.51 ± 9.76</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Reaction time for go trials, ms</td>
<td>—</td>
<td>300.84 ± 49.75</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Data shown as mean ± standard deviation unless otherwise specified. All tests were 2-tailed.
†Escitalopram 10–40 mg/d: 16; fluoxetine 40–100 mg/d: 15; paroxetine 37.5–87.5 mg/d: 9; sertraline 100–200 mg/d: 7; fluvoxamine 200–400 mg/d: 4.
‡Lorazepam-equivalent dose: 1.198 mg/d.
§Quetiapine 25 mg/d: 1; quetiapine 50 mg/d: 2; aripiprazole 5 mg/d: 1.
significantly decreased ICF compared with healthy controls, indicating reduced glutamatergic signalling in OCD. The effect size of decreased ICF was greater after excluding patients with comorbid depression ($\eta^2 = 0.154$). Conversely, Greenberg and colleagues$^{19}$ reported higher mean values for ICF in OCD than in controls, although this difference did not reach significance. Richter and colleagues$^{21}$ also showed that patients with OCD had greater ICF than healthy controls. As well, mixed findings have been observed for levels of cortical glutamate transmission in patients with OCD.$^{37,38}$ These inconsistent findings may be partially explained by the clinical heterogeneity of OCD, medication effects, limited statistical power and the different phase of treatment courses across studies. Although the findings remain inconclusive, the altered ICF observed in patients with OCD might reflect a disrupted neuroplasticity via altered glutamate-mediated excitatory neurotransmission in OCD. Evidence from TMS studies in depressive patients suggests that reduced facilitation and impaired glutamate-mediated neuroplasticity in response to paired associative stimulation are implicated in depression and cognitive dysfunction.$^{39,40}$ Further research is required to confirm the present findings and to gain a better understanding of alterations in cortical excitability and neuroplasticity in patients with OCD.

**Fig. 1:** Motor cortical excitability between patients with OCD and healthy controls. Graph showing variable means of bilateral RMT, CSP, ICF and SICI amplitudes in patients with OCD ($n = 51$) and controls ($n = 39$). Error bars represent ± 1 standard error of the mean. CSP = cortical silent period; ICF = intracortical facilitation; MEP = motor evoked potential; OCD = obsessive–compulsive disorder; RMT = resting motor threshold; SICI = short-interval intracortical inhibition.
Together, our findings of altered CSP and ICF provide further neurophysiological evidence for an imbalance in inhibitory and excitatory function in the cortical circuits of patients with OCD. This imbalance is also supported by substantial evidence from functional imaging studies, which have shown the involvement of CSTC circuits in OCD.4,41,42 Within the CSTC circuitry, the interactive loop between the GABAergic inhibitory and glutamatergic excitatory pathways is thought to be responsible for balancing neural tone.4,38 Genetic studies have also shown the involvement of genes related to GABA or glutamate in OCD pathogenesis.43–45 The imbalance of GABAergic and glutamatergic receptor–mediated neurotransmission in CSTC circuitry may contribute to the regional hyperactivity and lack of response inhibition seen in OCD, leading to obsessive–compulsive symptoms.

Limitations

Several limitations should be mentioned. First, although there was no difference in the degree of SICI between medicated and unmedicated patients in previous studies,19,33 we cannot exclude the influence of medications on motor cortex excitability. All patients in the present study were undergoing SRI treatment, and some were on concomitant benzodiazepines or antipsychotics. Alterations in TMS measures produced by medication have been reported to be complex depending on the individual drug and chronic use: a single dose of citalopram enhances CSP and SICI, whereas chronic paroxetine use does not alter either CSP or SICI but enhances ICF.46 In addition, benzodiazepines enhance SICI and reduce ICF, whereas dopamine antagonists, such as antipsychotics, increase ICF.47 For paired-pulse TMS measurements, medications influencing GABA, glutamate, serotonin or dopamine neurotransmission could affect GABAergic and glutamatergic receptor–mediated neurotransmission in CSTC circuitry and neuroplasticity.47,48 Second, the limited sample size may make it difficult to draw definitive conclusions, because low statistical power may lead to increased rates of false negatives and false positives. Third, since a role for the GABAergic and glutamatergic system has also been reported in the pathophysiology of depression, comorbid depression may have biased the present results. Previous TMS studies have shown that patients with depression had reduced excitability of both inhibitory and facilitatory inputs compared with controls.49,50 However, our results of shortened CSP and decreased ICF still remained significant after excluding individuals with comorbid depression. The correlation between CSP and response inhibition on the go/no-go task was also significant after adjusting for depressive symptoms. Therefore, we believe that the findings of altered cortical excitability may be specifically related to OCD, rather than depression. Fourth, this study is limited by the lack of neuroradiological imaging and of estimates of peripheral nerve excitability and central motor conductivity. Finally, we examined TMS indices of cortical excitability only in the motor cortex, which may not be relevant to brain regions critically involved in OCD. Measurements of cortical excitability in brain regions critically involved in OCD may be helpful for understanding the underlying pathophysiology of OCD.

Fig. 2: Correlation between CSP and mean reaction time for the correct go trials in the go/no-go task and age at onset in patients with OCD (n = 51). CSP = cortical silent period; OCD = obsessive–compulsive disorder.
Conclusion

The present study showed that patients with OCD had altered cortical excitability with shortened CSP and decreased ICF. In addition, the associations between CSP and response inhibition and onset age further suggest that cortical inhibition may be involved in the pathophysiology of OCD. Our findings support the role of altered cortical excitability in contributing to symptom formation in OCD. Further research in larger samples that include unmedicated patients is warranted to elucidate the pathophysiology of OCD.

Acknowledgements: This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2015R1D1A1A09058829).

Affiliations: From the Department of Psychiatry and Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, South Korea (Kang, C. Kim, S.J. Kim); the Department and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, South Korea (D.Y. Kim); and the Yonsei Phil Neuropsychiatric Clinic, Seoul, South Korea (Lee).

Competing interests: None declared.

Contributors: J.I. Kang wrote the article, which all authors reviewed. All authors acquired the data, which J.I. Kang and S.J. Kim analyzed. J.I. Kang, D.Y. Kim and S.J. Kim designed the study.

References


Randomized controlled trial of a gluten-free diet in patients with schizophrenia positive for antigliadin antibodies (AGA IgG): a pilot feasibility study

Deanna L. Kelly, PharmD; Haley K. Demyanovich, MPS; Katrina M. Rodriguez, MHS; Daniela Čiháková, PhD; Monica V. Talor, MSc; Robert P. McMahon, PhD; Charles M. Richardson, MD; Gopal Vyas, DO; Heather A. Adams, PsyD; Sharon M. August, MA; Alessio Fasano, MD; Nicola G. Cascella, MD; Stephanie M. Feldman, MSW; Fang Liu, MS; MacKenzie A. Sayer, BS; Megan M. Powell, BA; Heidi J. Wehring, PharmD; Robert W. Buchanan, MD; James M. Gold, PhD; William T. Carpenter, Jr., MD; William W. Eaton, PhD

Introduction

A connection between schizophrenia and inflammation is emerging in the literature, supported by many research approaches.1,2 Research also supports the idea that individual risk for schizophrenia is increased with prenatal infection exposure,3 and that a genetic contribution may also increase vulnerability to prenatal infection.4 Imaging studies have found increased microglia activation binding in people with schizophrenia,5–7 and genome-wide association studies have linked schizophrenia to the major histocompatibility locus.8–10 Epidemiological studies have shown higher rates of autoimmune disease in people with schizophrenia. For example, data from a Danish registry reveal a 50% greater likelihood of autoimmune disease in people with schizophrenia, and a 45% greater risk of schizophrenia in people with a history of autoimmune disease.11 Moreover, people with schizophrenia — even drug-naïve and first-episode patients — have been found to exhibit elevated levels of both central and peripheral cytokines.12

However, there are components of schizophrenia that remain unexplained by this inflammation hypothesis. For example, not all those with schizophrenia present with the same symptoms, not all experience inflammation, and research into inflammatory markers such as cytokines and chemokines yields mixed results. As with all major psychiatric disorders,
many genetic, environmental and biological factors contribute to a diagnosis. In this paper, we suggest that inflammation plays a considerable role for a specific subgroup of people with schizophrenia.13 This subgroup, once identified, may benefit from particular treatments that influence the immune system.

Evidence for a connection between the gut and the brain is emerging.14 Gut–brain connections may involve microbiota or gut-related immune reactions, including links to the gastrointestinal effects of wheat. The notion that schizophrenia may be connected to wheat consumption is not new,15 but it is still not yet well understood. The first epidemiological studies looking at the association between wheat and schizophrenia use data from World War II, when a positive correlation between wheat consumption and admissions for schizophrenia was first documented: in Scandinavia, as wheat consumption decreased, so did schizophrenia-related admissions; on the other hand, as wheat consumption rose in the United States, so did admissions for schizophrenia.16 Further, in populations that traditionally consume little to no grain, schizophrenia has been found to be almost nonexistent, but to exist after westernization.13,17,18

Several trials have explored the removal of wheat from the diets of people with schizophrenia, with mixed results: several showed improvement in psychiatric symptoms, but others did not.19–27 These findings were not well understood, because there were no biological markers to help define a subpopulation that might have benefited from the removal of wheat. Further, in the 1970s and 1980s the components of wheat were not well understood, and only a limited number of hypotheses existed as to why this connection may have been important.

Foods containing wheat, barley, rye and triticale (a wheat and rye cross) possess a protein composite called gluten. Gluten, the protein component from gluten that gives bread the ability to rise, has a low ratio of surface area to volume, making it difficult to digest. The autoimmune reaction to wheat is called celiac disease and has been recognized for about half a century, but more recently a new gluten-associated disorder has been described, distinct from celiac disease, that results from an innate immune reaction to gluten. This disorder has been called “gluten sensitivity,” but the understanding of what it constitutes remains somewhat controversial.28–34

About one-third of people with schizophrenia have antigluten antibodies of the immunoglobulin G (IgG) type,35 a rate about 3 times higher than seen in healthy controls.36–38 Also, the presence of AGA IgG antibodies in schizophrenia is related to a chronic inflammatory state associated with elevated peripheral cytokines (e.g., interleukin-1β and tumour necrosis factor α)39 and levels of neurochemicals in the anterior cingulate cortex thought to be associated with inflammation (as measured by magnetic resonance spectroscopy).40 Inflammation and immune activation with gluten sensitivity have been reported beyond schizophrenia as well.41

These findings in people with schizophrenia may be due to leakage in the blood–brain barrier. Severance and colleagues42 found a strong correlation between AGA IgG levels in the blood and cerebral spinal fluid in people with schizophrenia but not in healthy controls. Antibodies to tissue transglutaminase (tTG), indicative of celiac disease, are seen in only about 3%–5% of people with schizophrenia, representing a slightly higher risk than in the general population.43,44 Those with high AGA IgG levels represent a subgroup who may have gluten sensitivity.

We believe that schizophrenia is a heterogeneous multifactorial disease; the final symptoms may manifest similarly, but the underlying mechanisms producing psychiatric symptoms may differ by subgroup. Therefore, it may be key to use specific biomarkers to increase the efficacy signal of any intervention (in our case, a dietary one). Based on this premise, we decided to implement a gluten-free diet only in participants who showed elevated levels of AGA IgG. We believe that the reason for the mixed results in clinical trials during the 1970s and 1980s was the lack of ability to identify people at risk for this immune reaction and the resulting high inflammatory state.

Prior to this study, we conducted a 2-week gluten-free inpatient study for 2 people who had elevated AGA IgG and schizophrenia, and we noted robust symptom improvements, particularly in the domain of negative symptoms.45 The aims of this 5-week feasibility study were to create a gluten-free study design and successfully enrol more participants to examine effect sizes and plan a future trial of sufficient size with appropriate instrumentation. To our knowledge, this is the first double-blind, randomized, strictly controlled study exploring the removal of gluten from the diet in an inpatient setting that focused on a subgroup showing antibodies related to wheat. We hypothesized that we would see improvements in this group with effect sizes (Cohen d) of >0.5 in psychiatric symptoms.

Methods

Study procedures

Those who were eligible for screening had a diagnosis of schizophrenia or schizoaffective disorder, were not currently on a gluten-free diet and were between the ages of 18 and 64 years. Screening laboratory tests included AGA IgG, AGA IgA and tTG. If participants tested positive for tTG, they were excluded, and they and their clinical team were notified of their potential for celiac disease. Participants who tested positive for AGA IgG (>20 U) were eligible for study enrolment, which involved 5 weeks of randomized, double-blind treatment with a gluten-free or gluten-containing diet in an inpatient setting with strict dietary procedures. Participants who opted to enrol in the clinical trial were admitted to the research hospital unit and continued previous antipsychotic treatment. At the end of the 5-week double-blind trial, half of the discharged participants were randomly selected to continue a gluten-free diet in the community; they were called at week 4 for follow-up and returned at 8 weeks postdischarge for assessments and blood work.

Ethical approval

This clinical trial was conducted at the Maryland Psychiatric Research Center, University of Maryland School of Medicine. Screening took place at the center, its affiliate sites and Johns Hopkins University. All participants signed informed consent after passing the Evaluation to Sign Consent45 to ensure they were able to provide consent.
The study was conducted between 2014 and 2017 and was approved by the University of Maryland Baltimore institutional review board as the primary institutional review board (HP-00056339); the Johns Hopkins University and State of Maryland Department of Health institutional review boards officially relied on the University of Maryland Baltimore institutional review board for the conduct of this study. The study was reviewed annually by a data safety and monitoring board and was registered in ClinicalTrials.gov (NCT01927276).

Participants

Inclusion criteria
Women and men (ages 18–64 years) who met DSM-IV-TR criteria for schizophrenia or schizoaffective disorder were eligible for the study. All participants must have had positive results (> 20 U) on their AGA IgG screening. Participants must have been taking the same antipsychotic for at least 4 weeks prior to the study. As noted earlier, all participants were required to score at least a 10 out of 12 on the Evaluation to Sign Consent,\(^4\) which documented their capacity to provide informed consent.

Exclusion criteria
All those who participated in screening and tested positive for tTG were excluded so that their presumed celiac disease could be treated appropriately. Those who tested positive for AGA IgA but negative for AGA IgG were excluded. Also excluded were participants who were already on a gluten-free diet, were pregnant or lactating, had an organic brain disorder or intellectual disability, had a medical condition whose pathology or treatment could alter the presentation or treatment of schizophrenia or significantly increase the risk associated with the proposed treatment protocol, and who met DSM-IV criteria for alcohol or substance abuse (other than nicotine) within the last month. Participants were excluded if they had gluten ataxia, determined by a physician with the aid of the Brief Ataxia Rating Scale.\(^{47}\) To increase the possibility of detecting positive change, participants who had a Brief Psychiatric Rating Scale (BPRS)\(^{48}\) total score of 29 or lower (lower quartile) were excluded.

Screening assessments
Screening involved 1–2 visits and a variety of assessments to determine study eligibility. Participants were educated about the study and possible adverse effects or consequences of participation. Psychiatric diagnosis was confirmed by the Structured Clinical Interview for Diagnosis of DSM-IV (SCID).\(^{49}\) A medically accountable physician reviewed participants’ medical history and conducted a physical examination to confirm study eligibility. A standard blood chemistry panel, complete blood count, urinalysis and electrocardiography were also conducted. We also evaluated existing gastrointestinal disorders and dermatologic disorders, because gluten sensitivity may be related to both.\(^{51}\) For outpatients admitted to the unit and subsequently discharged following study participation, we worked closely with their community providers and supports to ensure full continuity of care.

Study assessments

Psychiatric symptoms
We measured psychiatric symptoms using the BPRS,\(^{48}\) the Scale for the Assessment of Negative Symptoms (SANS),\(^{49}\) the Calgary Depression Scale (CDS)\(^{50}\) and the Clinical Global Impression scale (CGI).\(^{51}\) We measured positive symptoms using the sum of the following BPRS items: conceptual disorganization, suspiciousness, hallucinatory behaviour and unusual thought content. We assessed negative symptoms using the total score of the SANS\(^{52}\) but subtracted the global items inappropriate affect, poverty of content of speech and attention to be consistent with other studies that measured negative symptoms in schizophrenia.\(^{53,54}\) All of these assessments are widely used in schizophrenia research and have good validity and reliability for use in this population. We administered the assessments each week during the 5 weeks of the study. All raters were trained and reliable, with an intraclass correlation coefficient of > 0.7.

We used the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) to assess cognitive function.\(^{55}\) This battery was developed specifically to be the gold standard for assessing cognition in schizophrenia.\(^{56}\) We also administered the Brief Assessment of Cognition in Schizophrenia Tower of London Test to assure adequate coverage of executive function.\(^{57}\) In recent multicentre clinical trials that our research team has participated in, individual tests of the MCCB have had intraclass correlation coefficients between 0.6 and 0.8, and the overall MCCB composite score has had an intraclass correlation coefficient of 0.9.

Adverse effect measures
We measured gastrointestinal effects using the Gastrointestinal Symptom Rating Scale (GSRS) at baseline and end point.\(^{57}\) This 15-item scale has 5 domains (reflux, constipation, diarrhea, pain and indigestion), and has been used in people with psychiatric symptoms.\(^{58}\) To assess medication adverse effects, we administered the Simpson–Angus ExtrapyramidalSymptom Rating Scale,\(^{59}\) the Barnes Akathisia Rating Scale\(^{60}\) and a 25-item adverse effect checklist each week.\(^{59,60}\)

Laboratory assessments
At screening and each week throughout the trial, we assessed weight and vital signs (heart rate, pulse, blood pressure). We drew a chemistry panel, including liver enzymes, lipids and fasting blood glucose, along with complete blood counts at baseline and end point, all analyzed by LabCorp. We also measured AGA IgG at baseline and end point in the Cˇiháková Laboratory at Johns Hopkins University using the Inova Diagnostics kit 708655. Units were determined by the manufacturer based on standard curve calculations. For this study, AGA IgG negative status was defined as < 20 U, and positive status as ≥ 20 U.
Randomization procedures and medications

Participants, researchers and clinical team members were blinded to intervention assignment. Participants received 10 g of either rice flour (treatment) or gluten flour (control; Bob’s Red Mill) mixed in a protein shake each afternoon. A hospital nurse mixed the blinded powder in a high-power blender with water, ice, protein powder and optional syrup of the participants’ choice to make a protein shake (Sunwarrior Plant-Based Protein). The research staff ensured that the entire shake was ingested. Each participant had a separate colour-coded container that was never shared to avoid cross-contamination. Treatments were assigned at random, using computer-generated permuted block randomization sequences with randomly varied block sizes to limit imbalance in the number of patients assigned to each group, while making it difficult for staff to predict what treatment patients were receiving. All participants received a gluten-free meal plan during the 5 weeks of the study. The hospital kitchen had fully functioning operating procedures for the creation of gluten-free meals, including 21 days of meals in a rotating schedule. During the study, participants attended weekly group sessions about gluten-free nutrition and diet counseling, and they had the opportunity to taste, shop for and prepare new gluten-free foods in this setting. We maintained a strict regimen and oversight for maintaining a gluten-free diet. All staff in the inpatient setting were required to ensure that all participants remained gluten-free. During the day, the staff:participant ratio was 1:1, and all nursing staff and aides were required to report if any deviation occurred. Each participant had an individual snack bucket with gluten-free snacks if needed, and no other free food was given out on the inpatient unit. All packaged food consumed was certified gluten-free. All participants received a gluten-free cookbook. After the end of their trial participation, study participants were discharged if they were considered stable according to the treating psychiatrist, and all were encouraged to follow a gluten-free diet.

The study intervention was added to participants’ ongoing antipsychotic regimen. Study physicians were instructed to avoid changing doses of other somatic and psychotropic medications during the study. Anticholinergic medications for extrapyramidal side effects (e.g., benzotropine and diphenhydramine), propranolol for akathisia and benzodiazepines for anxiety or agitation (e.g., lorazepam) could be prescribed as needed.

Statistical analysis

This was a pilot feasibility study funded by the National Institutes of Mental Health, aiming to ensure that we could develop and maintain blinding, recruit participants into the inpatient setting and assess a diet intervention in this population in a valid way. This study was not powered to find an effect, but SANS and BPRS scores were primary outcomes for symptom improvement, and the total GRS score was the primary outcome for gastrointestinal changes. We evaluated change in primary outcomes by calculating effect sizes using Cohen d that can be used to guide future confirmatory trials. An effect size is a quantitative measure of the magnitude of a difference — in this case change in findings from baseline to end point with a gluten-free diet relative to a gluten-containing diet. This was defined as treatment group mean – placebo group mean)/SDpooled.

\[ SD_{pooled} = \sqrt{\frac{(n_1-1)SD_1^2 + (n_2-1)SD_2^2}{n_1 + n_2 - 2}} \]

We also analyzed SANS results by week using least squares mean, the group mean calculated from an analysis of covariance model after controlling for covariates. We used the following model for the current study: change from baseline = baseline + treatment + week + treatment × week. This allowed us to control for baseline differences in the presentation of the data and was defined a priori. Statistical tests for analysis of covariance findings were not performed because of the small sample size and underpowering to show effect.

Results

Screening and participant information

During study recruitment, we screened 375 people with schizophrenia or schizoaffective disorder, of whom 100 (27%) screened positive for AGA IgG (> 20 U). Five of those were also positive for tTG (1.3%), indicating celiac disease; they were referred for treatment with their clinicians. In total, 64 people were interested in enrolling in the clinical trial, 26 of those met the eligibility criteria and 19 were randomized. Three were excluded after randomization because of low AGA IgG levels (< 20 U), leading to an enrolment total of 16 (9 in the gluten-containing group and 7 in the gluten-free group). Of those, 14 completed the study; 2 from the gluten-containing group discontinued the study early, 1 because of personal choice and 1 for housing reasons (Fig. 1).

The mean age of the 16 participants was 37.9 ± 13.2 years. During study recruitment, we screened 375 people with schizophrenia or schizoaffective disorder, of whom 100 (27%) screened positive for AGA IgG (> 20 U). Five of those were also positive for tTG (1.3%), indicating celiac disease; they were referred for treatment with their clinicians. In total, 64 people were interested in enrolling in the clinical trial, 26 of those met the eligibility criteria and 19 were randomized. Three were excluded after randomization because of low AGA IgG levels (< 20 U), leading to an enrolment total of 16 (9 in the gluten-containing group and 7 in the gluten-free group). Of those, 14 completed the study; 2 from the gluten-containing group discontinued the study early, 1 because of personal choice and 1 for housing reasons (Fig. 1).

The mean age of the 16 participants was 37.9 ± 13.2 years. The mean age of the 16 participants was 37.9 ± 13.2 years. The mean age of the 16 participants was 37.9 ± 13.2 years. The mean age of the 16 participants was 37.9 ± 13.2 years.

Table 1 lists the demographic and clinical characteristics of the groups. All participants maintained their doses of antipsychotics and other treatments during the 5-week study, and all were discharged on the regimen on which they had been initiated.

Psychiatric and neuropsychologic symptoms

We calculated the results of the primary analyses as effect sizes (Cohen d). Positive effect sizes indicated a greater average reduction in psychiatric scores with the gluten-containing diet; negative effect sizes indicated greater average reduction in psychiatric scores with the gluten-free diet.

Changes on the CGI showed a robust improvement in global symptoms (Cohen d = −0.75). We observed no improvement in the BPRS or CDS, which showed low effect sizes. However, notably for negative symptoms, we found an effect size of −0.53 for total SANS scores, suggesting a moderate improvement in negative symptoms (Fig. 2). We
found medium to large effect sizes in 2 of the SANS negative symptom domains: avolition (Cohen $d = -0.43$) and affective blunting (Cohen $d = -0.71$). We found small effect sizes for anhedonia (Cohen $d = -0.24$) and alogia (Cohen $d = -0.12$).

For the MCCB, a positive effect size score indicates greater average reduction in psychiatric scores with the gluten-free diet, because improvement in scores is in a positive direction (opposite of SANS, BPRS, CGI and CDS). The MCCB composite score effect size was not notable (Cohen $d = -0.18$). However, 2 of the domain scores had medium to large effect sizes favouring the gluten-free group: attention (Cohen $d = -0.66$) and verbal learning (Cohen $d = -0.37$).

**Gastrointestinal and other adverse effects**

All participants tolerated the diet well. The gluten-free group showed a robust improvement in total gastrointestinal symptoms, as measured by total GSRS score (Cohen $d = -0.81$), and medium to large improvements in domains of abdominal pain (Cohen $d = -0.86$), diarrhea (Cohen $d = -0.59$), constipation (Cohen $d = -0.40$) and indigestion (Cohen $d = -0.46$). The domain of reflux did not show any improvement (Cohen $d = 0.18$).

For adverse effects other than those on the adverse effect checklist, we found no differences between the 2 groups on any reported event, including dermatologic complaints. Three people in the gluten-free group had headaches, as did 2 in the gluten-containing group. Two people in the gluten-free group reported sedation, as did 1 person in the gluten-containing group. One person in each group reported dizziness and salivation. For extrapyramidal symptoms, we observed no notable changes in scores on the Simpson–Angus Extrapyramidal Symptom Rating Scale or the Barnes Akathisia Rating Scale.

**Table 1: Demographic and clinical information**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Gluten-containing diet ($n = 9$)</th>
<th>Gluten-free diet ($n = 7$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, yr</td>
<td>42.0 ± 14.6</td>
<td>32.5 ± 9.7</td>
</tr>
<tr>
<td>M/F, no. (%)</td>
<td>5/4 (56/44)</td>
<td>4/3 (57/43)</td>
</tr>
<tr>
<td>Age of onset, yr</td>
<td>16.9 ± 3.4</td>
<td>18.2 ± 2.4</td>
</tr>
<tr>
<td>Level of education, yr</td>
<td>11.8 ± 1.3</td>
<td>12.4 ± 2.1</td>
</tr>
<tr>
<td>Smoker Y/N, no. (%)</td>
<td>5/4 (56/44)</td>
<td>6/1 (86/14)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>28.5 ± 4.7</td>
<td>31.4 ± 8.9</td>
</tr>
<tr>
<td>Baseline AGA IgG, U</td>
<td>55.8 ± 28.6</td>
<td>43.8 ± 12.2</td>
</tr>
<tr>
<td>Baseline AGA IgA, U</td>
<td>23.6 ± 21.1</td>
<td>32.9 ± 28.3</td>
</tr>
<tr>
<td>Medications, no. (%)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA &lt; 5</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>SGA &lt; 5</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Clozapine &lt; 5</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>FGA + SGA &lt; 5</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>6 (67)</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Anticholinergic</td>
<td>7 (78)</td>
<td>6 (86)</td>
</tr>
<tr>
<td>Comorbid disorders, no. (%)‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal disorder</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Dermatologic disorder‡</td>
<td>&lt; 5</td>
<td>&lt; 5</td>
</tr>
</tbody>
</table>

FGA = first-generation antipsychotic; IgA = immunoglobulin A; IgG = immunoglobulin G; SGA = second-generation antipsychotic.

†Distribution of medications and comorbid disorders did not differ between groups.

‡Dermatologic disorders included eczema, urticaria and rash.

**Fig. 1:** Participant flow diagram. AGA = antigliadin antibodies; IgG = immunoglobulin G; tTG = tissue transglutaminase.
Laboratory measures

We observed no notable changes in findings for vital signs, electrocardiogram, complete blood count or chemistry panel, except for a small improvement in fasting blood glucose (Cohen $d = −0.36$), which favoured the gluten-free diet over the gluten-containing diet. Over the 5 weeks, AGA IgG levels decreased by 34% in the gluten-free group relative to 16% in the gluten-containing group (Cohen $d = −0.34$). As part of the protocol, the investigators had serum frozen for future analysis of inflammatory markers, kynurenine pathway metabolites$^{62}$ and potential measures of gut permeability (i.e., zonulin).$^{63}$

Eight-week follow-up

Eight randomly selected participants agreed to continue eating gluten-free for 8 weeks postdischarge. Five were from the gluten-containing group, and 3 from the gluten-free group. During follow-up, AGA IgG levels dropped 22% (mean ± standard deviation: 62.8 ± 36.9 at discharge to 49.0 ± 29.6 at 8 weeks) in participants from the gluten-containing group and 19% (27.8 ± 17.9 at discharge to 22.6 ± 13.5 at 8 weeks) in participants from the gluten-free group. The total SANS score remained relatively stable in each group over the 8 weeks: the gluten-containing group went from 25.0 ± 5.2 to 24.5 ± 4.1, and the gluten-free group maintained their benefits, going from 16.0 ± 8.2 to 14.0 ± 5.3.

Discussion

This is, to our knowledge, the first randomized, double-blind clinical trial of gluten withdrawal in people with schizophrenia that enrolled only participants likely to respond to a gluten-free diet (i.e., those who tested positive for AGA IgG). The design was rigorous, including strict rules and a consistent environment for gluten-free food preparation, as well as detailed blinding procedures for participants and staff. This study used a small sample to test feasibility, but we had robust findings in a few interesting domains. We noted an overall improvement on the CGL suggesting that overall clinical impressions during the study were better in the gluten-free group than in the gluten-containing group. We also saw moderate beneficial effects for negative symptoms. In our previous open-label study,$^{44}$ 1 person had a decrease of more than 19 points in SANS total score, leading us to anticipate the possibility of improvements in negative symptoms with a gluten-free diet in the present study. Because there are currently no consistently effective treatments for negative symptoms, the possible benefit of a gluten-free diet is of considerable potential clinical importance.

Other points worthy of discussion include the 27% rate of AGA IgG positivity in our study sample. Accumulating evidence suggests these antibodies are present in higher rates among people with schizophrenia than in healthy controls,$^{37,64}$ with about 30% of people with schizophrenia having AGA IgG seropositivity.$^{35,36}$ Additionally, it is unclear whether people with schizophrenia can maintain a gluten-free diet. Of the participants randomly selected to be followed postdischarge, improvements in AGA IgG levels and SANS total scores were maintained in the 3 participants who had been randomized to the gluten-free group during the 5-week inpatient study.

Limitations

A 5-week study may be insufficient to see amelioration of AGA IgG, psychiatric symptoms and cognitive function; longer-term...
studies may be needed to observe a full effect. Medications
developed to modify these processes may exist as alternatives to
strict gluten-free diets. We did not evaluate antibodies to
casein, which have also been studied in schizophrenia,63 and
we also did not evaluate foods containing fermentable oligo-
di/monosaccharides and polyols, which also may contribute
gastrointestinal symptoms or immune reactivity.64 We were
limited by the small sample and lack of power to show effect.
Our study purpose was to obtain the first go/no-go signal for a
full-size, double-blind, randomized clinical trial of this nature.

Conclusion

Our findings suggest that a subgroup of people with schizo-
phrenia may benefit from a gluten-free diet for global
improvement, negative symptoms and gastrointestinal ben-
fits. They also provide a better understanding of why some
groups of people with schizophrenia have high levels of
inflammation. If one-third of people with schizophrenia who
have AGA IgG were to benefit substantially from a gluten-free
diet, it could provide a new transformative treatment option
for an identifiable subpopulation of people with schizophre-
nia and be of enormous benefit to patients, families and soci-
ety. These results could also be used to encourage screening
for AGA IgG in people at high risk of schizophrenia or with
first-episode schizophrenia.

As a follow-up to this pilot study, we are conducting a larger
randomized double-blind confirmative clinical trial targeting
negative symptoms (NCT03183509), which will help confirm
the utility of gluten removal in schizophrenia patients who are
positive for AGA IgG if these findings are replicated. We are
also studying mechanisms related to zonulin measurement for
gut integrity,65 neuroimaging and inflammatory markers.

Acknowledgements: The authors thank the clinical and research team of the Treatment Research Program (TRP) at the Maryland Psychiatric Research Center (MPRC). The authors thank the nursing staff, who were instrumental in helping to maintain a gluten-free environment and helping with blood draws for screening, and the Outpatient Research Program (ORP) at the MPRC for their help in recruitment and screening. They also thank the many students, residents and fellows who spent countless hours with participants that helped with cooking classes and study integrity. They thank numerous students and train-
nees for their help over the years with the study, manuscript preparation and data analysis. Lastly, the authors thank the Research Pharmacist for his work in the preparation of the intervention materials. Prelimi-
nary data were presented at the Schizophrenia International Research Society Meeting in Florence, Italy, in April 2018.

Affiliations: From the Maryland Psychiatric Research Center (MPRC), School of Medicine, University of Maryland, College Park, MD (Kelly, McMahon, August, Feldman, Liu, Powell, Wehring, Buchanan, Gold, Carpenter); the Department of Orthopedics, School of Medicine, University of Maryland, College Park, MD (Demyanovich); the Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (Rodriguez, Eaton); the Department of Pathology, Division of Immunology, Im-
mune Disorders Laboratory, Johns Hopkins University, Baltimore, MD (Chihaková, Talor); the Spring Grove Hospital Center, Baltimore, MD (Richardson, Vyas, Adams); the Center for Celiac Research and Treatment, Massachusetts General Hospital, Boston, MA (Fasano); the Department of Psychiatry and Behavioral Sciences, Johns Hop-
kins University, Baltimore, MD (Cascella); the Department of Psych-
ology, Kent State University, Kent, OH (Sayer).

Funding: This study was funded by NIMH R34 (NIMH R34 MH100776; PIs: Eaton and Kelly). Preparation of this paper was supported by NIMH grant R01 MH113617; PIs: Kelly and Eaton. Clinical Trials.gov NCT19127276.

Competing interests: D. Kelly served as an advisor to Lundbeck and HLS Therapeutics. A. Fasano is the founder and a stock holder of Alba Therapeutics. R. Buchanan served on the advisory boards for Astellas Pharma, Avanir, Boehringer Ingelheim-RCV, ITI, Inc., Lund-
beck and Roche. He was a consultant for Takeda and Upsher-Smith Laboratories and on the DSMB for Pfizer. W. Carpenter has served as an advisor to Boehringer Ingelheim, Allergan, Health Analytics and Teva. All other authors have nothing to disclose.

Contributors: D. Kelly, H. Demyanovich, D. Chihaková, M. Talor, R. McMahon, A. Fasano, N. Cascella and W. Eaton conceived and de-
lished and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

References

1. Fineberg AM, Ellman LM. Inflammatory cytokines and neuro-
ological and neurocognitive alterations in the course of schizophre-
3. Flintham E, Sesh RK, Marttunen M, et al. Prenatal inflam-
4. Clarke MC, Tanskanen A, Huttunen M, et al. Evidence for an interaction between familial liability and prenatal exposure to infe-
6. Van Berckel BN, Bosnong MG, Boelaard R, et al. Microglia activation in recent-onset schizophrenia: a quantitative (R)-[11C]PK11195 posi-
8. Bergen SE, O’Dushlaine CT, Ripke S, et al. Genome-wide associa-
tion study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Mol Psychiatry 2012;17:880-6.
Kelly et al.

Neural response to emotional faces in monozygotic twins: association with familial risk of affective disorders

Iselin Meluken, MD; Ninja Ottesen, MD; Catherine Harmer, PhD, MA; Julian Macoveanu, MSc, PhD; Hartwig Siebner, MD, DMSc; Lars Kessing, MD, DMSc; Maj Vinberg, MD, DMSc, PhD; Kamilla Miskowiak, DPhil, DMSc

Background: Aberrant neural and cognitive response to emotional faces has been observed in people at familial risk of an affective disorder. In this functional MRI (fMRI) study of monozygotic twins, we explored neural correlates of the attentional avoidance of emotional faces that we had previously observed in high-risk versus affected twins, and whether an abnormal neural response to emotional faces represents a risk endophenotype. Methods: We recruited unaffected monozygotic twins with a co-twin history of mood episodes (high-risk), monozygotic twins with previous mood episodes (affected) and monozygotic twins with no personal or first-degree history of mood episodes (low-risk) between December 2014 and January 2017 based on a nationwide register linkage. Participants viewed fearful and happy faces while performing a gender discrimination task during fMRI and performed emotional faces dot-probe and facial expression recognition tasks outside the scanner. Results: A total of 129 monozygotic twins underwent whole-brain fMRI. High-risk twins (n = 38) displayed greater medial and superior prefrontal response to emotional faces than affected twins (n = 62). This greater activity correlated with stronger attentional avoidance of emotional faces in high-risk twins. In contrast, high-risk and affected twins showed no aberrant neural activity to emotional faces compared with low-risk twins (n = 29). Limitations: A limitation of this study was its cross-sectional design. Conclusion: Greater recruitment of the medial and superior prefrontal cortex during implicit emotion processing in high-risk versus affected twins may represent a compensatory or resilience mechanism. In contrast, aberrant neural response to emotional faces does not seem to be a risk endophenotype for affective disorders.

Introduction

Unipolar and bipolar disorders are prevalent psychiatric disorders with detrimental personal and societal consequences. However, the mechanisms that precede and prevent the onset of these affective disorders remain unclear. To elucidate such mechanisms and potentially improve opportunities for treatment and prevention, investigation of people at familial risk for an affective disorder is highly valuable. Because unaffected monozygotic twins have the same genetic makeup as their affected co-twins, studies of unaffected high-risk monozygotic co-twins provide a unique opportunity to investigate traits associated with familial risk.

In cross-sectional studies, unaffected people at familial risk for an affective disorder may display traits associated with that increased risk. At the same time, having withstood disease onset at the time of investigation, they may also display traits associated with resilience and compensatory adaptation. Although the onset of bipolar disorder peaks in adolescence and early adulthood, risk of disease onset has been shown to continue throughout adulthood. Consequently, firm conclusions about markers of risk, resilience and compensation can be drawn only from prospective high-risk studies in which comparisons are made between those who remain healthy at follow-up and those with the onset of disease. Nevertheless, cross-sectional studies that directly compare unaffected people at high risk with affected and low-risk control groups provide insight into potential markers of risk, resilience and compensation. Based on work by Wiggins, we define “risk endophenotypes” as traits shared by those who are affected or at high risk (versus those at low risk). These traits meet the endophenotype criterion of trait-related phenomena that are present in family members to a higher degree than in the general population. We define “markers of resilience” as traits shared by people at high risk and low risk relative to affected people.
individuals. These traits may represent mechanisms that promote mental health and characterize people with no past or present major psychiatric disorder. Finally, we define “markers of compensation” as traits that are specifically present in people at high risk relative to those who are affected and at low risk. These traits are likely to represent compensatory strategies implemented by these people to stay well despite their familial risk (Fig. 1).

Facial expressions are pivotal cues in the guidance of human behavior and are preferred stimuli when investigating neural correlates to basic emotion processing. Several functional MRI (fMRI) studies have observed that adult first-degree relatives and high-risk monozygotic and dizygotic twins display imbalances in corticobulbar response to emotional faces compared to low-risk groups. However, other studies have found no such differences in emotion-associated neural activity related to familial risk.

Accordingly, the aims of the present fMRI study were 2-fold: to investigate the neural correlates of the observed attentional avoidance of emotional faces in high-risk versus affected twins, including attentional avoidance of emotional faces. Attentional avoidance was measured as response latency when identifying probes were preceded by an emotional face in afaces dot-probe task. Accordingly, the aims of the present fMRI study were 2-fold: to investigate the neural correlates of the observed attentional avoidance of emotional faces in high-risk versus affected twins, and to investigate whether aberrant neural response to emotional faces represents a risk endophenotype that is present in high-risk twins and affected twins relative to low-risk twins, consistent with the idea that aberrant neural activity may be a more sensitive assay of abnormal brain functioning than behavioral measures. We therefore included behavioral tasks assessing different aspects of facial processing, such as recognition of and attention to emotional faces outside the scanner and an implicit facial-expression processing task previously used in similar independent twin samples to investigate neural correlates during fMRI.

**Methods**

**Participants**

A nationwide record linkage of the Danish twin registry and the Danish psychiatric central research register identified eligible monozygotic twins. In addition to monogygosity, eligibility criteria were age 18 to 50 years and a personal or co-twin history of a mood disorder (i.e., International Statistical Classification of Diseases and Related Health Problems, 10th revision, codes F30.0 to 34.0 and F38.0), or for low-risk twins neither a personal nor a co-twin history of a mood disorder. If only 1 twin from a twin pair was included, data from the Danish Central Research Register were used to determine risk status. Discordant status of twin pairs was assessed retrospectively with the Schedules for Clinical Assessment in Neuropsychiatry. To ensure familial low risk of major psychiatric disorders in unaffected twin pairs specifically, these pairs were excluded if they reported other first-degree relatives with an organic mental disorder, a schizophrenia-spectrum disorder or an affective disorder.

**Procedure and clinical assessment**

Participants were invited to attend a 1-day assessment. They underwent biological data sampling, clinical ratings of mood symptoms, a diagnostic interview, neurocognitive testing and fMRI scans (1 scan session lasted 1 hour and 2 minutes). We assessed lifetime diagnoses of psychiatric illness using the Schedules for Clinical Assessment in Neuropsychiatry. All twins were grouped according to personal and co-twin history of moderate to severe unipolar or bipolar disorder. Based on the International Statistical Classification of Diseases and Related Health Problems, 10th revision, those with a history of mixed states were included among participants with bipolar disorder. If only 1 twin from a twin pair was included, data from the Danish Central Research Register were used to determine risk status. Discordant status of twin pairs was defined as 1 twin with a lifetime history of moderate to severe depression or bipolar disorder and 1 twin without such a history, assessed retrospectively with the Schedules for Clinical Assessment in Neuropsychiatry.
Assessment in Neuropsychiatry interview. Objective rating instruments included the HDRS-17, the YMRS and the Danish Adult Reading Task to estimate premorbid verbal intelligence. All assessors were blinded for participants’ risk status. We obtained self-report ratings of mood symptoms and subjective state using the Major Depression Inventory,a visual analogue scale of current emotions and the State–Trait Anxiety Inventory form Y. We used the 10-item Edinburgh Inventory to assess handedness.

All participants gave informed consent to the study, conducted according to the Helsinki declaration. The study was approved by the local ethics committee (H-3–2014–003) and the Danish data protection agency (2014–331–0751).

Implicit emotional face processing during fMRI

We assessed neural response to happy and fearful faces with a block design paradigm. We presented 4 blocks of happy and fearful faces interleaved for 25 s. Each block consisted of 10 faces, starting with 5 female and ending with 5 male, all taken from the Nimstım Face Stimulus Set. Each face was presented for 200 ms, with an interstimulus interval of 2300 ms. A baseline fixation cross was presented between blocks for 20 s. Participants were instructed to categorize faces as male or female as quickly and correctly as possible by pressing 1 of 2 buttons. Participants’ responses were used to calculate mean reaction times and accuracy.

Behavioural tasks outside the scanner

We assessed attention to and recognition of emotional faces using the faces dot-probe and facial recognition tasks from the Oxford Emotional Test Battery.

In the faces dot-probe task, pairs of faces were presented horizontally, either unmasked with a duration of 100 ms or masked with a duration of 17 ms. One of the faces was replaced by 2 dots presented either vertically (·) or horizontally (· ·). Each face pair consisted of the same person with an emotional and a neutral expression, or with 2 neutral expressions. Participants were instructed to indicate the orientation of the dots as quickly and accurately as possible.

In the facial recognition task, faces expressing 1 of the 6 basic emotions — anger, disgust, fear, happiness, sadness and surprise — were displayed for 500 ms morphed at 10% intensity levels between a neutral face (0%) and a full emotional face (100%). Pictures of emotional faces were taken from Ekman and Friesen. Participants were instructed to categorize faces as happy or male as quickly and accurately as possible.

In the facial recognition task, faces expressing 1 of the 6 basic emotions — anger, disgust, fear, happiness, sadness and surprise — were displayed for 500 ms morphed at 10% intensity levels between a neutral face (0%) and a full emotional face (100%). Pictures of emotional faces were taken from Ekman and Friesen. Participants were instructed to categorize faces as happy or male as quickly and accurately as possible.

MRI data acquisition

All MRI scans were acquired at the Danish Research Centre for Magnetic Resonance at Copenhagen University Hospital Hvidovre using a 3 T Siemens Verio scanner and a 32-channel head array receive coil. During emotional faces processing, we acquired 140 volumes of $T_1^*$-weighted echo planar imaging with parallel imaging (GRAPPA) and a whole-brain field of view (acceleration factor = 2, field of view 192 mm², matrix size 64 × 64, axial imaging plane, slice thickness 3 mm, 42 slices, interleaved upwards acquisition order, echo time 30 ms, repetition time 2320 ms, flip angle 80°). We acquired $T_1^*$-weighted images for participant alignment using an MPRAGE sequence (field of view 230 mm², slice thickness 0.9 mm, 224 slices, repetition time 1900 ms, echo time 2320 ms, flip angle = 9°). We recorded participants’ pulse and respiration during the scan.

Analysis of fMRI data

Preprocessing and single-subject (first-level) analysis

We conducted analyses using FEAT version 5.0.9, part of the FMRIB Software Library. Standard preprocessing steps included nonbrain removal, linear and nonlinear registration to structural space, normalization to the Montreal Neurological Institute (MNI) standard space, motion correction and spatial smoothing using a Gaussian kernel of 5 mm full width at half maximum. We corrected for geometric distortions based on an acquired B0 field map. All participants’ registration and unwarping results were visually controlled. Additionally, before high-pass temporal filtering (cutoff 90 s), we carried out an independent component analysis (ICA)-based strategy for the automatic removal of motion artifacts. Finally, we performed manual ICA-based denoising to remove components resulting from acquisition artifacts.

The first-level general linear model included 4 regressors of interest, modelling response to fearful male, fearful female, happy male and happy female faces. The dependent variable was the estimated $\beta$ weights of the general linear model. We included a regressor of no interest modelling participants’ failure to indicate sex if the number of missing answers exceeded 2 standard deviations of the mean. We modelled all regressors by convolving each with a double-$\gamma$ hemodynamic response function. We also included temporal derivatives of task regressors in the model as covariates of no interest to model slice-timing effects. We performed physiologic noise modelling cardiac and respiratory noise, creating 16 additional regressors to model out these effects. A priori contrasts of interest were an emotional face-processing response of happy and fearful relative to (>) baseline, and negative and positive valence-specific responses of fearful > happy and happy > fearful.

Group (second-level) analysis

For our first objective, we compared affected and high-risk twins (i.e., independent variables) using a 2-sample t test and an intrapair analysis of complete discordant twin pairs using a paired-sample t test. In the paired-sample t test, we included discordant twin pairs without fMRI data from both twins.

For our second objective, we compared high-risk twins with affected and low-risk twins (i.e., independent factor levels) using analysis of variance. The 3 contrasts of interest were dependent variables in all models. We conducted group-level analyses using permutation inference with
permutation analysis of linear models, restricting permutation to within and between twin pairs and between single twins. We performed analyses with and without adjustment for depressive symptoms (i.e., HDRS-17 score) within 2 volumes of interest (VOI), and an exploratory analysis across the whole brain. First, we used a mask consisting of the anterior cingulate cortex and the paracingulate cortex based on findings of negative functional connectivity between these regions and the amygdala in our previous twin studies and on the key role of these areas in conflict monitoring, attribution of mental states and implicit emotion regulation. Second, we used a larger mask to explore areas shown to be involved in emotional face processing in affective disorders, consisting of the superior, inferior and middle frontal gyri; the frontal pole; the frontal medial cortex; the anterior cingulate cortex; the paracingulate gyrus; the temporal and occipital fusiform cortex; the subcallosal cortex and frontal orbital cortex; the insular cortex; the parahippocampal gyrus; and the bilateral hippocampus and amygdala. We made the VOI masks using the Harvard–Oxford cortical and subcortical atlases implemented in FSLview, thresholded at 20%. We determined cluster-wise thresholding using the threshold-free cluster enhancement method and we accepted family wise error–corrected p values of < 0.05 as significant. We reported peak activation of significant clusters using MNI coordinates and cerebral regions with corresponding Brodmann areas, identified through Talairach conversion of MNI coordinates with GingerALE and a standard anatomic atlas. We extracted the percent blood-oxygenation level–dependent signal change from significant clusters using the featquery tool for illustrative purposes and post hoc correlation analysis. We investigated correlation of the extracted percent signal change in each risk group separately with attentional avoidance of emotional faces and with depressive symptoms (i.e., HDRS-17 scores). We calculated avoidance of emotional faces as a mean vigilance score of unmasked fearful and masked happy conditions from the faces dot-probe task, based on observations from the full monozygotic sample. Additionally for the discordant twin pairs, we investigated correlation of the extracted percent signal change in high-risk twins with age at illness onset for affected twins and discordant time. We calculated discordant time as the time passed between illness onset for the affected twin and the assessment of the high-risk twin.

Functional connectivity analysis
We conducted psychophysiological interaction analysis to assess functional connectivity with functional clusters in the left and/or right structural amygdala as seed regions. We defined the functional clusters as areas in the left or right amygdala that displayed significant activation to emotional faces across all participants, derived from a 1-sample t test. This resulted in a functional cluster in the left amygdala (98 voxels; MNI coordinates: x, y, z = −20, −2, −14; peak p < 0.001), and no significant activation in the right amygdala. We entered the seed region time-course from the functional cluster in the left amygdala in a psychophysiological interaction model that included all original regressors and 4 additional psychophysiological interaction regressors. We investigated the interaction of left amygdala time-course and response to emotional faces > baseline, fearful faces > baseline and happy faces > baseline.

Analysis of behavioural data
We examined sex discrimination during scanning, vigilance to fear and happiness, and recognition of facial expressions in general and of happiness and fear specifically (i.e., dependent variables) using mixed-model analysis of variance, with group as fixed factors and twin pairs and participants as random factors. In the analysis of the 10 intensity levels of happy and fearful faces (i.e., dichotomous variables), we used logistic regression techniques with nested random effects for twin and participant. We modelled emotional expressions in the facial recognition and sex discrimination tasks as within-group factors in repeated-measures models. The 2 high-threshold models were applied to obtain a measure of discrimination accuracy for facial expressions corrected for response tendency. Conducted data analyses in SAS 9.4 (SAS Institute Inc.).

Results

Participants
Of the 204 participants included in the overall study of putative endophenotypes for affective disorders, a subsample of 134 twins underwent fMRI (high-risk: n = 38; affected: n = 66; low-risk: n = 30). The reasons for not scanning 70 participants were as follows: target fMRI sample size reached (n = 16); group size for affected twins with unipolar disorder reached to ensure a balanced sample (n = 29); participants declined (n = 16); exclusion because of metal in the body or head trauma (n = 4); or other reasons (n = 5). Two participants did not complete the scan session because of claustrophobia or excessive noise. Of the remaining 132 participants, 1 was excluded because of sex discrimination accuracy of 3 standard deviations below the mean, and 2 were excluded because of technical issues. The analyses of fMRI data included 129 monozygotic twins (high-risk: n = 38; affected: n = 62; low-risk: n = 29).

Demographic and clinical characteristics are presented in Table 1. Among the 38 high-risk twins, 28 (74%) had a co-twin diagnosed with unipolar disorder and 10 (26%) had a co-twin diagnosed with bipolar disorder. The sample included 13 concordant (bipolar disorder/bipolar disorder: n = 3; unipolar disorder/unipolar disorder: n = 5; bipolar disorder/unipolar disorder: n = 5), 22 discordant (high-risk/unipolar disorder: n = 15; high-risk/bipolar disorder: n = 7) and 11 low-risk complete twin pairs, as well as 37 single twins (co-twin included without fMRI data: n = 27; co-twin not included: n = 10). The 3 groups were well balanced with respect to age, sex, years of education, premorbid IQ and handedness (Table 1). As expected, affected twins scored higher on subsyndromal depressive symptoms and trait and state anxiety, and lower on happiness than high- and low-risk twins (p ≤ 0.01).
Table 1: Demographic and clinical comparison of affected, high-risk and low-risk monozygotic twins (n = 129)

<table>
<thead>
<tr>
<th>Variable*</th>
<th>Affected twins (n = 62)</th>
<th>High-risk twins (n = 38)</th>
<th>Low-risk twins (n = 29)</th>
<th>F value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>37.5 (35.2–39.8)</td>
<td>36.6 (33.7–39.5)</td>
<td>37.4 (34.1–40.8)</td>
<td>F_{1,120} = 0.12</td>
<td>0.77</td>
</tr>
<tr>
<td>Education, yr</td>
<td>14.8 (14.0–15.6)</td>
<td>15.7 (14.7–16.7)</td>
<td>15.5 (14.3–16.7)</td>
<td>F_{1,120} = 1.01</td>
<td>0.36</td>
</tr>
<tr>
<td>Premorbid IQ*</td>
<td>113.5 (111.9–115.1)</td>
<td>112.0 (109.8–114.2)</td>
<td>110.6 (105.6–115.6)</td>
<td>F_{1,120} = 1.05</td>
<td>0.34</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>43 (69)</td>
<td>26 (68)</td>
<td>21 (72)</td>
<td>F = 0.76</td>
<td>0.68</td>
</tr>
<tr>
<td>Left-handed (LQ &lt; 0), n (%)</td>
<td>10 (16)</td>
<td>9 (24)</td>
<td>&lt; 5</td>
<td>F = 3.7</td>
<td>0.16</td>
</tr>
<tr>
<td>Bipolar I disorder, n (%)</td>
<td>19 (31)</td>
<td>NA</td>
<td>NA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Bipolar II disorder, n (%)</td>
<td>5 (8)</td>
<td>NA</td>
<td>NA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Unipolar disorder, n (%)</td>
<td>38 (61)</td>
<td>NA</td>
<td>NA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>No. of episodes</td>
<td>4.3 (3.3–5.4)</td>
<td>NA</td>
<td>NA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Age at onset, yr</td>
<td>24.2 (22.4–26.0)</td>
<td>NA</td>
<td>NA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Medications, n (%)</td>
<td>Antidepressant 25 (40)</td>
<td>&lt; 5</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lithium</td>
<td>12 (19)</td>
<td>0</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Anticonvulsant</td>
<td>12 (19)</td>
<td>0</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>12 (19)</td>
<td>0</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Comorbid disorders, n (%)</td>
<td>Anxiety disorder 5 (8)</td>
<td>5 (13)</td>
<td>&lt; 5</td>
<td>F = 1.00</td>
<td>0.37</td>
</tr>
<tr>
<td>Prior substance abuse</td>
<td>&lt; 5</td>
<td>0</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Other†</td>
<td>5 (8)</td>
<td>0</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Clinical assessment scores</td>
<td>Hamilton Depression Rating Scale</td>
<td>4.3 (3.6–5.0)</td>
<td>2.6 (1.7–3.5)</td>
<td>1.9 (0.9–2.9)</td>
<td>F_{1,118} = 9.22</td>
</tr>
<tr>
<td></td>
<td>Young Mania Rating Scale</td>
<td>1.9 (1.4–2.3)</td>
<td>1.5 (0.9–2.1)</td>
<td>1.4 (0.8–2.1)</td>
<td>F_{1,118} = 0.71</td>
</tr>
<tr>
<td></td>
<td>Major Depression Inventory</td>
<td>7.7 (6.3–9.5)</td>
<td>4.6 (3.5–6.1)</td>
<td>3.5 (2.0–6.3)</td>
<td>F_{1,118} = 6.75</td>
</tr>
<tr>
<td></td>
<td>State–Trait Anxiety Inventory, state</td>
<td>30.2 (28.5–31.9)</td>
<td>27.2 (25.3–27.5)</td>
<td>26.3 (24.2–28.6)</td>
<td>F_{1,118} = 4.43</td>
</tr>
<tr>
<td></td>
<td>State–Trait Anxiety Inventory, trait</td>
<td>39.7 (38.0–41.5)</td>
<td>33.0 (31.2–35.0)</td>
<td>34.1 (32.0–36.4)</td>
<td>F_{1,118} = 14.88</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, happiness</td>
<td>5.0 (4.6–5.4)</td>
<td>6.1 (5.6–6.6)</td>
<td>5.8 (5.2–6.4)</td>
<td>F_{1,118} = 6.15</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, sadness</td>
<td>1.2 (0.9–1.6)</td>
<td>0.7 (0.3–1.2)</td>
<td>0.8 (0.3–1.3)</td>
<td>F_{1,118} = 1.73</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, vigilance</td>
<td>3.7 (3.0–4.4)</td>
<td>3.7 (2.8–4.6)</td>
<td>5.6 (3.5–7.7)</td>
<td>F_{1,118} = 1.93</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, anxiety</td>
<td>0.8 (0.5–1.0)</td>
<td>0.5 (0.1–0.8)</td>
<td>0.4 (0.0–0.8)</td>
<td>F_{1,118} = 1.26</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, dizziness</td>
<td>0.6 (0.4–0.9)</td>
<td>0.6 (0.3–0.9)</td>
<td>0.3 (0.1–0.6)</td>
<td>F_{1,118} = 1.43</td>
</tr>
<tr>
<td></td>
<td>Visual analogue scale, nausea</td>
<td>0.4 (0.2–0.6)</td>
<td>0.4 (0.2–0.7)</td>
<td>0.3 (0.0–0.6)</td>
<td>F_{1,118} = 0.36</td>
</tr>
</tbody>
</table>

LQ = lateral quotient; NA = not applicable.

*Unless otherwise indicated, descriptive and clinical variables are presented as estimated group means with confidence intervals calculated using a mixed-model procedure, accounting for dependence within twin pairs. Group comparisons of affected, high-risk and low-risk twins are reported with F values and p values. Counts of fewer than 5 were suppressed owing to data privacy guidelines.

† Measured by the Danish adult reading task; 9 participants with dyslexia were excluded.

‡Attention-deficit/hyperactivity disorder, eating disorder, adjustment disorder.

fMRI results

Main effect of task across participants
Table 2 presents the statistically significant main effects of task and group comparisons of high-risk, affected and low-risk monozygotic twins in VOI and whole-brain analysis. Emotional faces activated the superior frontal gyrus (SGF) in the medial prefrontal cortex (mPFC) VOI, as well as the fusiform gyri, middle frontal gyri, right SGF, right insular cortex and left cingulate gyrus in the emotional face-processing network VOI. In the whole-brain analysis, emotional faces activated cortical and subcortical areas involved in vision, motor and emotion processing, consistent with our group’s previous work using the same paradigm on a similar population (Fig. 2).14,15 We observed no statistically significant main effects of the fear > happy or happy > fear contrasts or amygdala seed-based functional connectivity.

Comparison of high-risk twins with affected twins
Within the mPFC VOI, high-risk twins showed increased activity to emotional faces in the bilateral medial frontal gyrus (mPFC, BA-8) and the SFG (BA-10; Table 2, Fig. 3). Within the mPFC cluster, the mean vigilance score to emotional faces correlated negatively with activity in high-risk twins (r = −0.4, p = 0.04, Fig. 3). However, an outlier analysis removing 2 participants with values more extreme than 3 standard deviations of the mean rendered this correlation nonsignificant (r = −0.3, p = 0.14). We found no significant correlation between vigilance scores to emotional faces in affected and low-risk groups and activity in the mPFC or with activity in the SFG across all groups (Table 1). Activity in these clusters also showed no correlation with depressive symptoms (Table 1). In a post hoc exploratory analysis adjusted for depressive symptoms, the difference in mPFC activity between high-risk and affected twins was reduced to
Table 2: Task and group comparisons of affected, high-risk and low-risk monozygotic twins in volume-of-interest and whole-brain analysis: statistically significant main effects*

<table>
<thead>
<tr>
<th>Region</th>
<th>Brodmann area</th>
<th>Right/left</th>
<th>MNI coordinates,† x, y, z</th>
<th>Voxels</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main effect across participants to emotional faces &gt; baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial prefrontal cortex volume of interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>6</td>
<td>Left</td>
<td>−6, 10, 44</td>
<td>878</td>
<td>0.0002</td>
</tr>
<tr>
<td>Emotional face processing network volume of interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>11</td>
<td>Right</td>
<td>22, 42, −18</td>
<td>34</td>
<td>0.037</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>45</td>
<td>Right</td>
<td>42, 26, 18</td>
<td>1864</td>
<td>0.0002</td>
</tr>
<tr>
<td>Insula</td>
<td>13</td>
<td>Right</td>
<td>32, 24, 4</td>
<td>81</td>
<td>0.02</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>32</td>
<td>Left</td>
<td>−6, 12, 44</td>
<td>797</td>
<td>0.0002</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>6</td>
<td>Left</td>
<td>−40, 4, 24</td>
<td>1685</td>
<td>0.0002</td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>37</td>
<td>Right</td>
<td>44, −44, −28</td>
<td>1146</td>
<td>0.0002</td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>37</td>
<td>Left</td>
<td>−38, −44, −28</td>
<td>1040</td>
<td>0.0002</td>
</tr>
<tr>
<td><strong>Whole brain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insula</td>
<td>13</td>
<td>Right</td>
<td>32, 24, 4</td>
<td>277</td>
<td>0.020</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>44</td>
<td>Right</td>
<td>46, 8, 22</td>
<td>3209</td>
<td>0.001</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>6</td>
<td>Right</td>
<td>38, 8, 64</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Brain stem</td>
<td>NA</td>
<td>Right</td>
<td>10, −64, −48</td>
<td>75865</td>
<td>0.0002</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>32</td>
<td>Right</td>
<td>8, 20, 36</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>44</td>
<td>Left</td>
<td>−34, 4, 22</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>24</td>
<td>Left</td>
<td>−4, 2, 46</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Sulcus cinguli</td>
<td>24</td>
<td>Left</td>
<td>−26, −8, 42</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Brain stem</td>
<td>NA</td>
<td>Right</td>
<td>4, −20, −20</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Transverse temporal gyri</td>
<td>41</td>
<td>Left</td>
<td>−48, −24, 16</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>High-risk twins &gt; affected twins to emotional faces &gt; baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial prefrontal cortex volume of interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td>8</td>
<td>Right</td>
<td>8, 42, 26</td>
<td>202</td>
<td>0.027</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>10</td>
<td>Right</td>
<td>10, 54, 4</td>
<td>72</td>
<td>0.039</td>
</tr>
</tbody>
</table>

FWE = family-wise error; MNI = Montreal Neurological Institute.
*Significant findings for main effects of task and group × task interactions are presented by cluster size and peak cluster localization, with corresponding peak p values. Results are derived from permutation methods that allowed us to model the dependence structure of twin pairs. To define clusters, we used the threshold-free cluster enhancement method, and found significant results by thresholding FWE-corrected images at $p_{\text{FWE}} = 0.05$. Results from the 2 volumes of interest used as small-volume correction and across the whole brain are presented.
†MNI coordinates refer to local maxima within cluster.

Fig. 2: Main effect of emotional faces relative to baseline across participants ($n = 129$) revealed robust activation in areas involved in emotional face processing. The bar represents values from a 1-sample t test.
Face processing in monozygotic high-risk twins

We found no differences between high-risk and affected twins in terms of neural response to emotional faces > baseline or to fearful versus happy faces within the emotional face processing network VOI or across the whole brain. We also found no group differences in functional connectivity from the left amygdala during face processing or in the intra-pair comparisons of discordant twin pairs (Appendix 1, available at jpn.ca/170246-a1).

Comparison of high-risk, affected and low-risk twins
We found no group differences in neural response to emotional faces > baseline, fearful > happy or happy > fearful faces with or without adjustment for depressive symptoms. Moreover, functional connectivity from the left amygdala during face processing did not differ between groups.

Behavioural results

Main effect of task across participants
Participants displayed good task compliance during scanning, as reflected by high accuracy of sex discrimination (group mean accuracy ≥ 97%). Participants were generally slower to discriminate when faces displayed fear than happiness (F(1,126) = 4.6; p = 0.04). In the behavioural tasks outside the scanner, participants displayed subliminal avoidance of happy faces (t(114) = –4.2; p < 0.001) and a positive bias in face recognition, as reflected by greater accuracy (F(1,114) = 107.3; p < 0.001) and lower speed (F(1,113) = 175.3; p < 0.001) for happy versus fearful faces. Behavioural data are presented in Appendix 1, Table S1.

Comparison of high-risk twins with affected twins
The 2 groups did not differ in terms of accuracy (p = 0.69) or speed (p = 0.54) during sex discrimination. High-risk twins displayed attentional avoidance of consciously processed fearful faces relative to affected twins (t(111) = –2.0; p = 0.045), and we observed a trend-level difference in subconsciously processed happy faces (t(111) = –1.9; p = 0.06). We observed no group differences in attention to consciously processed fearful faces.

Fig. 3: Results from volume-of-interest analysis, including medial areas with increased activity in high-risk twins versus affected twins. (A) The medial prefrontal cortex volume of interest, including areas involved in implicit emotion regulation and appraisal of affective stimuli previously shown to be aberrant in monozygotic twins using the same facial-processing paradigm (marked in medium grey). Also displayed is the area with significant main effect of task across participants within this mask (marked in dark grey). Finally, this panel also displays the 2 significant clusters in the medial and superior prefrontal cortex with increased activation to emotional faces over baseline (marked in white) in high-risk twins (n = 38) compared with affected twins (n = 62). (B) Blood-oxygenation level–dependent activity presented as mean percent signal change to emotional faces over baseline in high-risk twins (n = 38) and affected twins (n = 62). Percent signal change is presented as group mean with standard error of the mean computed by a mixed model, with twin pairs as random factors and group as fixed factors. Error bars represent standard error of the mean. (C) Significant correlation between the extracted mean percent signal change in response to emotional faces and a mean vigilance score of emotional faces in high-risk twins, with the corresponding Pearson coefficient and p value.
happy or subconsciously processed fearful faces, in facial expression recognition of fear versus happy and across the 6 emotional expressions or in recognition of the 10 intensity levels of happy and fearful faces (Table 1).

**Comparison of high-risk, affected and low-risk twins**

We observed no group differences in neither speed or accuracy of sex discrimination, vigilance to emotional faces, recognition of fearful versus happy faces, general facial expression recognition (across all 6 emotions) or recognition of fearful and happy faces across the 10 intensity levels (Table 1).

**Discussion**

Using whole-brain fMRI, we investigated the neural correlates of previously observed attentional avoidance of emotional faces in high-risk relative to affected monozygotic twins with a mood disorder. We also investigated whether aberrant neural activity represented a risk endophenotype that was present across high-risk and affected twins relative to low-risk twins. The results revealed that attentional avoidance of emotional faces in high-risk versus affected twins was accompanied by heightened response to emotional faces in the medial and superior PFC. This greater PFC activity correlated with more attentional avoidance in high-risk twins. In contrast, we observed no evidence for aberrant neural response to emotional faces representing a risk endophenotype, because we observed no shared imbalances in high-risk and affected twins versus low-risk twins. Notably, the difference in PFC response to emotional faces between high-risk and affected twins was reduced to a trend level when adjusting for subsyndromal depressive symptoms. We observed no difference in PFC response to emotional faces between medicated and nonmedicated affected twins, suggesting that medication did not confound our findings.

The greater recruitment of the medial and superior PFC in high-risk relative to affected twins is noteworthy, given that these regions are involved in implicit emotion regulation and conflict monitoring. Accordingly, the task requirement to focus on nonemotional aspects of faces may have introduced greater conflict monitoring and/or stronger implicit down-regulation of reactivity to the task-irrelevant emotional aspects of faces in these high-risk twins. Given this, the greater activity in the medial and superior PFC and its correlation with more avoidance of emotional faces could indicate that high-risk twins compensate for their familial risk. This interpretation is consistent with evidence for negative functional connectivity between the mPFC/dorsal anterior cingulate cortex and the amygdala from previous studies of high-risk groups, although see also the studies by Wiggins and Amico. Additionally, a 20-year prospective study of people at high versus low familial risk of unipolar disorder indicated that greater activity in overlapping regions, including the dorsal anterior cingulate cortex and SFG during an attention interference task was a marker of resilience that protected high-risk individuals from illness onset. Notably, however, the between-group difference in PFC response to emotional faces in the present study was reduced to a trend after adjustment for subsyndromal symptoms. Because affected twins had more subsyndromal depressive symptoms than high-risk twins, it is unclear whether the increased PFC activity to emotional faces contributed to fewer subsyndromal depressive symptoms in high-risk twins (in line with a compensatory role of the PFC) or if lower PFC activity contributed to more subsyndromal depressive symptoms in affected twins (representing a scar of illness). Notably, we found main and group effects for neural activity to emotional faces in general, but no specific effects for happy or fearful faces. In line with other reports of group differences in neural response to emotional faces in general, this suggests that at-risk individuals process emotional faces differently, regardless of emotional expression.

The absence of shared imbalances in high-risk and affected versus low-risk groups in terms of neural response to emotional faces was unexpected. Indeed, this lack of evidence for aberrant neural response to emotional faces as a risk endophenotype contrasts with previous observations of decreased activity in the dorsal PFC and increased mPFC activity to emotional faces in high-risk versus low-risk unipolar disorder or bipolar disorder groups. Additionally, unaffected people at high-risk and patients with bipolar disorder have been found to display similar exaggerated mPFC activity to emotional faces compared with low-risk groups. Further, in an independent sample of monozygotic twins, we found heightened medial and superior PFC activity in monozygotic twins at high versus low familial risk of unipolar disorder. Possible reasons for the different findings are a larger sample size in the high-risk group (n = 13), lower mean age (37 yr in the present study v. 47 yr), participant characteristics (mixed unipolar disorder and bipolar disorder in the present study v. unipolar disorder only) and analysis methods. Nonetheless, the notion of aberrant neural response to emotional faces representing a risk endophenotype for affective disorders is challenged by the negative findings of this and other studies.

**Limitations**

Several limitations in addition to the cross-sectional design should be mentioned. First, the use of psychotropic medication in affected twins might have influenced their neural response, because selective serotonin reuptake inhibitors have been shown to reduce aberrant limbic activity. Nevertheless, we found no differences in medial or superior PFC activity to emotional faces in exploratory post hoc comparisons of medicated and nonmedicated affected twins. Second, the use of a black screen with a fixation cross as baseline (as opposed to a neutral face stimulus) may be considered a limitation because this hinders disentangling of face-related and emotion-specific neural activity. However, the use of a neutral face baseline has also been criticized for not being perceived as neutral, but as negative instead, which could introduce a bias in the results. Third, the low-risk group was relatively small (n = 29) compared with the other groups, and the actual sample size was reduced in statistical inference because of dependant observation within twin pairs. Specifically, permutation was...
restricted in 2 levels: within twin pairs as well as between single twins and twin pairs. Fourth, we chose not to control for sex because there was an equal distribution of men and women across the groups, and the effects of risk were thus unlikely to be influenced by sex. Nevertheless, there is some evidence that men and women display different neural activity patterns during face processing. In fact, post hoc analysis comparing percent signal change between men and women within the significant mPFC clusters revealed greater blood-oxygenation level-dependent signal in one of these clusters in men compared with women (mPFC; \( p = 0.04 \)).

Fifth, the differential mPFC response in high-risk versus affected twins was reduced to a trend-level difference in a post hoc analysis that controlled for subsyndromal mood symptoms. This could suggest that the difference between these groups was because of the slightly higher subsyndromal symptom levels in the affected twins than in the high-risk twins (with average HDRS-17 scores of 4.3 v. 2.6, respectively) or to a reduction in the statistical power in these fMRI analyses that already involved control for any behavioural differences and physiologic noise. Finally, we focused on the common neural mechanisms of affective disorder (i.e., both unipolar disorder and bipolar disorder) based on shared symptomatology and genetic underpinnings. In post hoc analyses comparing percent signal change in the significant clusters in medial areas, we found no difference between participants affected with unipolar disorder versus bipolar disorder (mPFC: \( p = 0.25 \); SFG: \( p = 0.48 \)). However, studies investigating disorder-specific markers are also warranted to help increase diagnostic precision.

**Conclusion**

The greater recruitment of medial and superior PFC during implicit emotion processing in high-risk relative to affected twins — and its correlation with more attentional avoidance and emotional faces processing: a voxel-based meta-analysis of 105 participants with bipolar disorder. *Front Hum Neurosci* 2011;5:184.


**Competing interests:** C. Harmer has received consultancy fees from P0vital Ltd, Lundbeck, Servier and Eli-Lilly, and is a company director of Oxford Psychologists Ltd. She has also received grant income from GSK, UCBI, Janssen and the contributions to K. Miskowiak from Lundbeck, Servier and AstraZeneca. H. Siebner discloses honoraria as journal editor from Elsevier Publishers and book editor from Springer Publishing, as well as honoraria as speaker from Genzyme and MerckSeren, and grant support from Biogen-idec within the last 3 years. M. Vinberg discloses consultancy fees from Lundbeck within the last 3 years. L. Kessing has been a consultant for Sunovion within the last 3 years. K. Miskowiak reports having received consultancy fees from Lundbeck and Allergan in the past 3 years. No other competing interests declared.

**Contributors:** I. Meluken, N. Ottesen, L. Kessing, M. Vinberg and K. Miskowiak designed the study. I. Meluken, N. Ottesen, J. Macoveanu, H. Siebner, M. Vinberg and K. Miskowiak acquired the data, which I. Meluken, C. Harmer, J. Macoveanu, H. Siebner, L. Kessing, M. Vinberg and K. Miskowiak analyzed. I. Meluken, M. Vinberg and K. Miskowiak wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

**References**


**Management of sexual adverse effects induced by atypical antipsychotic medication**

Laura Downing, BN, MD; David D. Kim, MSc; Ric M. Procyshyn, PharmD, PhD; Philip Tibbo, MD

A 28-year-old man recently diagnosed with schizophrenia was discharged from hospital on long-acting injectable risperidone (37.5 mg given every 2 weeks). At an outpatient visit 2 months later, his psychotic symptoms were well controlled, but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.

As risperidone was thought to be the cause of the patient’s sexual dysfunction, the dose was reduced to 25 mg every 2 weeks. Unfortunately, his auditory hallucinations re-emerged, and there was no appreciable change in his sexual function. The serum prolactin level obtained at this time was 180 ng/mL (reference range for men: 3–15 ng/mL). After discussion with the patient, we decided to cross-titrate to 15 mg/d of aripiprazole over a period of 4 weeks, which reduced his prolactin to 7 ng/mL. On this regimen, the psychosis stabilized but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.

As risperidone was thought to be the cause of the patient’s sexual dysfunction, the dose was reduced to 25 mg every 2 weeks. Unfortunately, his auditory hallucinations re-emerged, and there was no appreciable change in his sexual function. The serum prolactin level obtained at this time was 180 ng/mL (reference range for men: 3–15 ng/mL). After discussion with the patient, we decided to cross-titrate to 15 mg/d of aripiprazole over a period of 4 weeks, which reduced his prolactin to 7 ng/mL. On this regimen, the psychosis stabilized but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.

As risperidone was thought to be the cause of the patient’s sexual dysfunction, the dose was reduced to 25 mg every 2 weeks. Unfortunately, his auditory hallucinations re-emerged, and there was no appreciable change in his sexual function. The serum prolactin level obtained at this time was 180 ng/mL (reference range for men: 3–15 ng/mL). After discussion with the patient, we decided to cross-titrate to 15 mg/d of aripiprazole over a period of 4 weeks, which reduced his prolactin to 7 ng/mL. On this regimen, the psychosis stabilized but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.

As risperidone was thought to be the cause of the patient’s sexual dysfunction, the dose was reduced to 25 mg every 2 weeks. Unfortunately, his auditory hallucinations re-emerged, and there was no appreciable change in his sexual function. The serum prolactin level obtained at this time was 180 ng/mL (reference range for men: 3–15 ng/mL). After discussion with the patient, we decided to cross-titrate to 15 mg/d of aripiprazole over a period of 4 weeks, which reduced his prolactin to 7 ng/mL. On this regimen, the psychosis stabilized but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.

As risperidone was thought to be the cause of the patient’s sexual dysfunction, the dose was reduced to 25 mg every 2 weeks. Unfortunately, his auditory hallucinations re-emerged, and there was no appreciable change in his sexual function. The serum prolactin level obtained at this time was 180 ng/mL (reference range for men: 3–15 ng/mL). After discussion with the patient, we decided to cross-titrate to 15 mg/d of aripiprazole over a period of 4 weeks, which reduced his prolactin to 7 ng/mL. On this regimen, the psychosis stabilized but he reported reduced libido and anorgasmia. The occurrence of these symptoms coincided with risperidone initiation. The patient had no previous history of sexual dysfunction, was not taking any other medications and denied any forms of substance use.
References

Research Paper

Impact of white matter hyperintensity location on depressive symptoms in memory-clinic patients: a lesion–symptom mapping study

Anna E. Leeuwis, MSc; Nick A. Weaver, MD; J. Matthijs Biesbroek, MD, PhD; Lieza G. Exalto, MD, PhD; Hugo J. Kuijf, PhD; Astrid M. Hooghiemstra, PhD; Niels D. Prins, MD, PhD; Philip Scheltens, MD, PhD; Frederik Barkhof, MD, PhD; Wiesje M. van der Flier, PhD; Geert Jan Biessels, MD, PhD; on behalf of the TRACE-VCI study group

Background: We investigated the association between white matter hyperintensity location and depressive symptoms in a memory-clinic population using lesion–symptom mapping. Methods: We included 680 patients with vascular brain injury from the TRACE-VCI cohort (mean age ± standard deviation: 67 ± 8 years; 52% female): 168 patients with subjective cognitive decline, 164 with mild cognitive impairment and 348 with dementia. We assessed depressive symptoms using the Geriatric Depression Scale. We applied assumption-free voxel-based lesion–symptom mapping, adjusted for age, sex, total white matter hyperintensity volume and multiple testing. Next, we applied exploratory region-of-interest linear regression analyses of major white matter tracts, with additional adjustment for diagnosis. Results: Voxel-based lesion–symptom mapping identified voxel clusters related to the Geriatric Depression Scale in the left corticospinal tract. Region-of-interest analyses showed no relation between white matter hyperintensity volume and the Geriatric Depression Scale, but revealed an interaction with diagnosis in the forceps minor, where larger regional white matter hyperintensity volume was associated with more depressive symptoms in subjective cognitive decline (β = 0.26, p < 0.05), but not in mild cognitive impairment or dementia. Limitations: We observed a lack of convergence of findings between voxel-based lesion–symptom mapping and region-of-interest analyses, which may have been due to small effect sizes and limited lesion coverage despite the large sample size. This warrants replication of our findings and further investigation in other cohorts. Conclusion: This lesion–symptom mapping study in depressive symptoms indicates the corticospinal tract and forceps minor as strategic tracts in which white matter hyperintensity is associated with depressive symptoms in memory-clinic patients with vascular brain injury. The impact of white matter hyperintensity on depressive symptoms is modest, but it appears to depend on the location of white matter hyperintensity and disease severity.

Introduction

Late-life depression is highly prevalent in older people and in patients with cognitive impairment or dementia. It has been associated with vascular dementia, stroke and white matter hyperintensity. This link between vascular disease and late-life depression has led to the “vascular depression hypothesis.” The clinical profile of vascular depression includes loss of interest and motivation, executive dysfunction and psychomotor retardation. The vascular depression hypothesis has been investigated intensively in population-based studies. Late-life depression is consistently associated with severity of white matter hyperintensity (i.e., visual rating scores; Fazekas or Scheltens scale) and larger total white matter hyperintensity volumes in healthy elderly people. Meanwhile, studies in memory-clinic populations are scarce, even though this could be a clinically important population, considering their generally higher vascular lesion burden and frequent occurrence of depressive symptoms. We recently showed that in memory-clinic patients with Alzheimer disease, the severity of white matter hyperintensity (measured using the Fazekas scale) was not related to depressive symptoms. However, we found a borderline significantly increased propensity for depressive symptoms in patients with subjective cognitive decline and white matter hyperintensity. Apart from the severity of white matter hyperintensity, recent studies suggest that specific locations of white matter hyperintensity could predispose people for depressive symptoms.

Correspondence to: A.E. Leeuwis, Alzheimer Amsterdam and Department of Neurology, Amsterdam Neuroscience, Amsterdam UMC, P.O. Box 7057, 1007 MB Amsterdam, the Netherlands; a.leeuwis@vumc.nl

Submitted Aug. 8, 2018; Revised Dec. 6, 2018; Accepted Jan. 9, 2019; Published online Apr. 25, 2019

DOI: 10.1503/jpn.180136

© 2019 Joule Inc. or its licensors

J Psychiatry Neurosci 2019;44(4) E1
The LADIS study found that deep white matter hyperintensities specifically located in the frontal and temporal locations were associated with depressive symptoms in nondisabled older people.11 Frontal white matter hyperintensities have been associated with higher depression scores on a questionnaire in patients with dementia.12 Furthermore, white matter hyperintensities in the prefrontal and temporal regions, and in specific white matter tracts such as the cingulum bundle, uncinate fasciculus and superior longitudinal fasciculus, have been associated with severity of depression in patients with major depression.13,14 These results suggest disruption of prefrontal–subcortical pathways in particular as an underlying mechanism of late-life depressive symptoms in elderly people.5 Identifying specific white matter tracts in which white matter hyperintensities have the most impact on depressive symptoms would improve our understanding of the consequences of cerebral vascular injury.

Lesion–symptom mapping is frequently used to investigate the relationship between lesion location and specific clinical symptoms in patients with vascular brain injury such as white matter hyperintensity, infarcts and lacunes. Most lesion–symptom mapping studies on white matter hyperintensity have focused on the association between white matter hyperintensity location and cognitive function,15–17 while psychological symptoms of subcortical vascular lesions, such as depression and anxiety, have not been addressed. In this first-ever lesion–symptom mapping study on depressive symptoms, we aimed to determine the extent to which specific white matter hyperintensity locations contribute to depressive symptoms in memory-clinic patients with vascular brain injury on MRI, and to identify strategic white matter tracts in which white matter hyperintensities affect depressive symptoms.

Methods

The TRACE-VCI (Utrecht–Amsterdam clinical features and prognosis in vascular cognitive impairment) study is a prospective observational follow-up study of 860 consecutive memory-clinic patients from Dutch outpatient clinics at 2 university hospitals: VU University Medical Centre and University Medical Centre Utrecht.18 All patients visited the memory clinic between September 2009 and December 2013 and underwent a 1-day standardized dementia screening process that included a medical history, physical and neurologic examinations, screening laboratory tests, an MRI scan of the brain and a neuropsychological assessment. Patients with cognitive complaints and any burden of vascular brain injury on MRI were prospectively included. Further inclusion and exclusion criteria are described in detail elsewhere.18 Patients were divided into 3 categories related to the extent of their cognitive impairment: dementia, mild cognitive impairment and subjective cognitive decline. Patients with evidence of co-occurring neurodegenerative disease or depression were accepted because these are common comorbid etiologies in patients with vascular cognitive impairment. We excluded patients with a nonvascular or nondegenerative primary cause of cognitive impairment, such as a brain tumour, extensive traumatic head injury, substance or alcohol abuse, or multiple sclerosis. We also excluded patients with a primary psychiatric disease other than depression. The study was approved by the medical ethics committee of VU University Medical Centre and University Medical Centre Utrecht. We obtained written informed consent from participants (or their responsible guardians if they were incapable of consent) before conducting research-related procedures.

Participants

A flow chart of patient selection for the present study is presented in Figure 1. Of the 860 patients in TRACE-VCI, 37 were excluded during the vascular lesion segmentation process, mostly because the available MRI data were of insufficient quality, or because of technical errors during data processing. Next, 100 of the remaining patients were excluded based on the presence of nonlacunar infarcts or hemorrhages other than microbleeds on MRI, because such large lesions can result in the complete obliteration of white matter tracts and could have interfered with our analysis in which white matter hyperintensity volume with specific tracts is related to depressive symptoms at a group level. One additional patient was excluded because of failed lesion registration. Finally, 42 patients were excluded because a Geriatric Depression Scale (GDS) score was not available. This resulted in a study sample of 680 patients (168 subjective cognitive decline, 164 mild cognitive impairment and 348 dementia).

For all patients, we determined history of depression and use of antidepressant medication (e.g., selective serotonin reuptake inhibitors, tricyclic antidepressants, monoamine oxidase inhibitors) based on self-reported medical history and medication use. We identified the presence of hypertension based on self-reported medical history, medication use or newly diagnosed hypertension, defined as a blood pressure of 140/90 mm Hg or higher and measured with a sphygmomanometer. We identified hypercholesterolemia based on self-reported medical history or medication use. We identified diabetes mellitus based on self-reported medical history, medication use or newly diagnosed diabetes mellitus, defined as a nonfasting blood glucose of 11.1 mmol/L or greater, or a glycosylated hemoglobin higher than 48 mmol/mol (or ≥ 6.5%). We defined obesity as a body mass index higher than 30 kg/m².

Evaluation of depressive symptoms

We assessed depressive symptoms using the 15-item self-reported GDS,19 which has a maximum score of 15; higher scores indicate the presence of depressive symptoms. The GDS is used frequently in clinical practice and research, and is a valid and reliable screening instrument for depressive symptoms in older adults.19 In our study, the GDS was verbally administered to patients by a neuropsychologist. We classified patients as having depressive symptoms if their GDS score was 5 or higher. In our analyses we used the continuous GDS score, because it offers the highest power to detect associations.
Alzheimer disease biomarkers

Cerebrospinal fluid markers β-amyloid1–42 (Aβ1–42) and total tau were available for 446 patients. We assessed cerebrospinal fluid biomarkers using Sandwich enzyme-linked immunosorbent assays (Fujirebio). Assays were considered positive for Alzheimer disease when the tau/Aβ1–42 ratio was > 0.52. In the patients selected for this study, cerebrospinal fluid biomarkers were measured only in those included at the VU University Medical Centre, as a standard procedure of the memory clinic.

MRI protocol

Brain MRI scans were performed on 1.5 T (n = 39) or 3.0 T (n = 641) MRI scanners. Scans were acquired on GE (n = 527, 77.5%) or Philips (n = 153, 22.5%) MRI scanners using a standardized protocol that included 3D T1-weighted, T2-weighted, T2*-weighted/susceptibility-weighted imaging (SWI) and T2 fluid-attenuated inversion recovery (FLAIR) sequences. For some patients, 3D T1 and/or FLAIR sequences were not available, so 2D T1 or FLAIR sequences were used instead. Slice thickness, voxel size and other details for each scanner type are described in detail in Appendix 1, Table S1, available at jpn.ca/180136-a1.

Lesion segmentation

We rated vascular brain injury in accordance with the internationally established STRIVE criteria, which provide neuroimaging standards for classification of cerebral small vessel disease. Ratings were performed by or under the supervision of a neuroradiologist. Lesion segmentation was performed on T2 FLAIR images, using the T1 modality as a reference for proper lesion classification. Automated white matter hyperintensity segmentation was performed using the k nearest neighbour (kNN-TTPs) method. This method showed no systematic errors across the different

---

**Fig. 1:** Flow chart for patient selection. GDS = Geriatric Depression Scale.
MRI scanners. The resulting white matter hyperintensity lesion maps underwent a visual check for accuracy by 2 independent raters. Subsequent manual corrections were required in 6 participants (0.9%) because of segmentation inaccuracies (i.e., missed white matter hyperintensity or incorrect or incomplete white matter hyperintensity segmentation). These corrections were performed by a single rater. In addition, we determined the presence of other lesion types: lacunes were defined as sharply demarcated deep lesions with cerebrospinal fluid–like signal on all sequences; microbleeds were defined as small dot-like hypointense lesions on T2*-weighted or SWI images. We performed manual segmentation of these lesions using software developed in house based in MeVisLab (MeVis Medical Solutions AG).24,25

**Generation of lesion maps**

All lesion maps were transformed to the T1 1 mm Montreal Neurological Institute (MNI-152) brain template,26 using an image registration pipeline developed in house that applies the elastix toolbox.27 This standardized pipeline has been recently developed and will soon be made publicly available at www.metavcimap.org. The registration procedure consisted of linear registration followed by nonlinear registration. As an intermediate step, we performed registration to an age-specific MRI template,28 which has been shown to result in more successful registration of brains from patients with severe atrophy. These registration steps were combined into a single step, through which the original lesion maps were registered directly to the MNI-152 space to prevent intermediate interpolations and improve registration accuracy. Quality checks of all registration results were performed by 1 rater (N.A.W.), who compared the lesion location in the MNI-152 space with the original scans. One patient (0.14%) was excluded because of unsuccessful lesion registration.

**Statistical analysis**

We used PASW Statistics 25.0 for Mac (SPSS Inc.) to conduct statistical analyses. We performed analyses of variance and Pearson χ² tests to compare groups when appropriate.

We applied 2 independent hypothesis-free analysis methods to identify white matter hyperintensity locations associated with depressive symptoms: voxel-based lesion–symptom mapping (VLSM), which analyzed the relation between presence of white matter hyperintensities and depressive symptoms for each voxel in the brain;29 and exploratory region-of-interest (ROI) analyses, which analyzed the impact of lesion volume in predefined white matter tracts on depressive symptoms.

**Voxel-based lesion–symptom mapping**

We performed VLSM using nonparametric mapping software (NPM, version May 2016; settings: univariate analysis, Brunner-Munzel test),29 which is suitable for non-normally distributed data. To ensure that our analyses were not biased by voxels that are only rarely affected, we set a minimum number of patients with a lesion in a particular voxel and included only voxels that were affected by white matter hyperintensity in at least 14 participants (2%).30 We performed VLSM analyses using a z-score of the GDS as a measure for depressive symptoms after individualized correction for age and sex using linear regression. We repeated the analyses after additional correction for normalized total white matter hyperintensity volume (i.e., calculated from lesion maps after transformation to MNI-152 space). We applied false discovery rate control (q < 0.05) to correct for multiple testing. We performed VLSM in the whole group and then stratified it for syndrome diagnosis.

**Region-of-interest analysis**

We created ROIs using the John Hopkins University diffusion tensor imaging–based white matter atlas31 with a probability threshold of 10%. We calculated regional white matter hyperintensity volumes in millilitres for each patient for 20 white matter tracts. Next, we merged bilateral white matter tracts to create a single ROI by combining the volumes. The GDS was standardized into a z-score. White matter hyperintensity volumes in the resulting 11 ROIs were added as independent variables to linear regression models, which included age, sex and memory-clinic centre of inclusion as covariates (Model 1). When we found a significant association, we repeated the analysis with additional adjustment for normalized total white matter hyperintensity volume (Model 2). We also performed extra analyses with adjustments for antidepressant medication and MRI field strength and vendor (Model 3). To investigate whether associations with the ROIs differed according to diagnostic group (subjective cognitive decline, mild cognitive impairment or dementia), we included interaction terms (dummy diagnosis × ROI) in the model. When we found an interaction between diagnosis and the ROI (p < 0.10), we stratified the results for syndrome diagnosis and displayed the standardized βs (β) for each diagnostic group separately. When no significant interaction was found, the interaction term was removed from the model and the overall β was reported.

Finally, we performed an additional linear regression analysis in a subgroup of patients with cerebrospinal fluid biomarkers (n = 446) to determine whether the impact of white matter hyperintensity location was influenced by co-occurring Alzheimer disease pathology. To investigate whether associations differed among patients with positive versus negative cerebrospinal fluid biomarkers, we used interaction terms (amyloid status × ROI).

**Results**

Demographic data and MRI measures are summarized in Table 1. We noted no differences between the original TRACE-VCI cohort and the present study sample (data not shown). Patients with subjective cognitive decline were younger than patients with mild cognitive impairment or dementia. Patients with dementia had lower scores on the GDS than patients with subjective cognitive decline (dementia
White matter hyperintensities and depressive symptoms

Table 1: Demographics of the study population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total sample (n = 680)</th>
<th>SCD (n = 168)</th>
<th>MCI (n = 164)</th>
<th>Dementia (n = 348)</th>
<th>Statistical test*</th>
<th>p value</th>
<th>Post hoc differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr, mean ± SD</td>
<td>67.1 ± 8.2</td>
<td>62.9 ± 7.5</td>
<td>68.2 ± 8.5</td>
<td>68.6 ± 7.7</td>
<td>F = 31.27</td>
<td>&lt; 0.001</td>
<td>SCD &lt; MCI = D</td>
</tr>
<tr>
<td>Education level, mean ± SD†</td>
<td>4.9 ± 1.3</td>
<td>5.1 ± 1.4</td>
<td>5.2 ± 1.2</td>
<td>4.7 ± 1.3</td>
<td>F = 9.47</td>
<td>&lt; 0.001</td>
<td>SCD = MCI &gt; D</td>
</tr>
<tr>
<td>MMSE score, mean ± SD</td>
<td>24.3 ± 4.8</td>
<td>27.7 ± 2.2</td>
<td>26.5 ± 2.3</td>
<td>21.7 ± 5</td>
<td>F = 169.15</td>
<td>&lt; 0.001</td>
<td>SCD &gt; MCI &gt; D</td>
</tr>
<tr>
<td>GDS score, mean ± SD</td>
<td>3.7 ± 3</td>
<td>4.6 ± 3.5</td>
<td>3.7 ± 2.8</td>
<td>3.2 ± 2.7</td>
<td>F = 12.66</td>
<td>&lt; 0.001</td>
<td>SCD &gt; MCI = D</td>
</tr>
<tr>
<td>Female, n (%)§</td>
<td>320 (47)</td>
<td>83 (49)</td>
<td>72 (44)</td>
<td>165 (47)</td>
<td>χ² = 1.045</td>
<td>0.593</td>
<td>NS</td>
</tr>
<tr>
<td>Depressive symptoms, n (%)§‡</td>
<td>200 (29)</td>
<td>72 (42)</td>
<td>51 (31)</td>
<td>77 (22)</td>
<td>χ² = 23.750</td>
<td>&lt; 0.001</td>
<td>SCD &gt; MCI &gt; D</td>
</tr>
<tr>
<td>Alzheimer disease biomarkers, n (%)§</td>
<td>446 (66)</td>
<td>112 (67)</td>
<td>100 (61)</td>
<td>234 (67)</td>
<td>χ² = 2.054</td>
<td>0.358</td>
<td>NS</td>
</tr>
<tr>
<td>Available</td>
<td>577 (84)</td>
<td>133 (79)</td>
<td>141 (86)</td>
<td>303 (87)</td>
<td>χ² = 5.717</td>
<td>0.057</td>
<td>NS</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>287 (42)</td>
<td>66 (39)</td>
<td>81 (49)</td>
<td>140 (40)</td>
<td>χ² = 4.615</td>
<td>0.100</td>
<td>SCD &lt; MCI &gt; D</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>123 (18)</td>
<td>23 (13)</td>
<td>39 (23)</td>
<td>61 (17)</td>
<td>χ² = 5.653</td>
<td>0.054</td>
<td>NS</td>
</tr>
<tr>
<td>Obesity (BMI ≥ 30 kg/m²)</td>
<td>144 (21)</td>
<td>43 (25)</td>
<td>34 (20)</td>
<td>67 (19)</td>
<td>χ² = 3.687</td>
<td>0.450</td>
<td>NS</td>
</tr>
</tbody>
</table>
| Education level was classified according to the Verhage system, ranging from 1 (low education) to 7 (highly educated).
| Presence of depressive symptoms indicates a score of ≥ 5 on the Geriatric Depression Scale.
| Available as cerebrospinal fluid total tau/β A1–42 (abnormal when > 0.5221).
| BMI = body mass index; D = dementia; GDS = Geriatric Depression Scale; IQR = interquartile range; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; NS = not significant; SCD = subjective cognitive decline; SD = standard deviation; WMH = white matter hyperintensity.

3.2 ± 2.7 v. subjective cognitive decline 4.6 ± 3.5; p < 0.001).

Patients with mild cognitive impairment or dementia used antidepressant medications less often than patients with subjective cognitive decline (dementia 12% and mild cognitive impairment 11% v. subjective cognitive decline 20%; p < 0.05). The total white matter hyperintensity volume was highest in patients with dementia and mild cognitive impairment, versus patients with subjective cognitive decline: dementia (median [IQR]) 11.5 [22.8] and mild cognitive impairment 11.6 [18.7] v. subjective cognitive decline 5.0 [10.7], p < 0.001).

Voxel-based lesion–symptom mapping

We used VLSM as an assumption-free method of investigating whether the presence of white matter hyperintensity in specific voxels of the brain was significantly associated with depressive symptoms on the GDS, independent of total white matter hyperintensity volume. The distribution of white matter hyperintensity is illustrated by the lesion prevalence map in Figure 2A. White matter hyperintensities showed a symmetric distribution, and were most prevalent in the periventricular and frontoparietal regions.

The results of the VLSM analysis are shown in Figure 2B. We found voxels with a significant association between the presence of white matter hyperintensities and depressive symptoms after correction for age, sex, total white matter hyperintensity volume and multiple testing. These significant voxels were almost exclusively located in the corticospinal tract, near the superior longitudinal fasciculus and the temporal part of the superior longitudinal fasciculus. The exact number of significant voxels in each white matter tract is provided in Table 2.

Subsequent stratification for syndrome diagnosis showed no significant voxels for any subgroup.

Region-of-interest analyses

We used ROI analyses to determine whether white matter hyperintensity volumes in predefined white matter tracts were associated with depressive symptoms. Table 3 shows the association between total and regional white matter hyperintensity volumes and depressive symptoms.

Neither total white matter hyperintensity volume nor regional white matter hyperintensity volume in specific tracts were related to depressive symptoms. We did find interactions
between syndrome diagnosis and regional white matter hyperintensity volume in the forceps minor, the anterior thalamic radiation, the inferior fronto-occipital fasciculus and the inferior longitudinal fasciculus, suggesting that the association between depressive symptoms and regional white matter hyperintensity volume in these regions is different for subjective cognitive decline, mild cognitive impairment and dementia. Subsequent stratification for syndrome diagnosis showed that in patients with subjective cognitive decline, regional white matter hyperintensity in the forceps minor was associated with more depressive symptoms ($\beta = 0.16; p < 0.05$). Additional adjustment for normalized total white matter hyperintensity volume resulted in a slightly stronger association ($\beta = 0.26; p < 0.05$). We found no significant associations for mild cognitive impairment or dementia. Finally, we performed analyses with additional adjustment for antidepressant medication and MRI field strength and vendor, and the results were unchanged (data not shown).

Exploratory region-of-interest analyses in a subgroup of patients with cerebrospinal fluid biomarkers

We performed exploratory analyses in a subgroup of patients with cerebrospinal fluid biomarkers ($n = 446$; Appendix 1, Table S1). We found no significant interactions between amyloid status and white matter hyperintensity volume in relation to depressive symptoms in any region.

Discussion

This lesion–symptom mapping study on depressive symptoms indicates the corticospinal tract and forceps minor as strategic white matter tracts in which white matter hyperintensities were associated with depressive symptoms in a memory-clinic cohort of patients with vascular brain injury. The overall impact of white matter hyperintensities on these symptoms was modest, but their location appeared to

Fig. 2: Voxel-based lesion–symptom mapping: lesion prevalence map and results. (A) Voxel-wise lesion prevalence of white matter hyperintensities in the study population, projected on the Montreal Neurological Institute 152 T1 template. A minimum threshold of 14 participants with damage in a given voxel was applied; z-coordinates: −5, 5, 15, 25, 35. Voxel-based lesion–symptom mapping results for the Geriatric Depression Scale score, shown in (B) axial, (C) sagittal and (D) coronal planes. Significant voxels after correction for multiple comparisons, age, sex and normalized total white matter hyperintensity volume are shown in red (settings: Brunner–Munzel test; FDR $q < 0.05$). Significant voxels were located in the corticospinal tract. Regions of interest were derived from the Johns Hopkins University diffusion tensor imaging atlas with a probability threshold of 10%. The corticospinal tract derived from the Johns Hopkins University atlas is shown in blue; the voxels included in the voxel-based lesion–symptom mapping analysis (i.e., damaged in ≥ 14 participants) are shown in yellow. Coordinates: sagittal: $x = −25$; coronal: $y = −32$; axial: $z = 33, 38$. CST = corticospinal tract; FDR = false detection rate; WMH = white matter hyperintensity.
be particularly important in patients with subjective cognitive decline.

The analyses used in this study (VLSM and ROI linear regression) resulted in different strategic white matter hyperintensity locations. We detected an association between regional white matter hyperintensity in the corticospinal tract and depressive symptoms only at the voxel level, and the number of significant voxels was limited (only 15 of 5975).

With the ROI analyses at the regional level, we found no congruent correlation with the corticospinal tract but identified a modest association between the forceps minor and GDS only in the subgroup with subjective cognitive decline. The statistical power (due to more rigorous correction for multiple testing) for the VLSM analyses might have been insufficient. However, a main advantage of VLSM is its very high spatial resolution. Still, our results for the forceps minor were consistent with previous findings on the role of white matter hyperintensities in frontal and temporal locations in depression.31 The lack of convergence of findings from the VLSM and ROI analyses may be because effect sizes are quite small. Moreover, the exploratory nature of the ROI analysis warrants replication of our findings and further investigation in other large memory-clinic cohorts with optimal lesion coverage.

The vascular depression hypothesis proposes that white matter hyperintensities caused by cerebrovascular disease disrupt the frontostriatal–subcortical circuits and predispose people for late-life depression.32 Previous studies on white matter pathways and depressive symptoms have primarily employed diffusion tensor imaging. Most studies examined patients with major depressive disorder or late-life depression. A recent review on white matter alterations in emotional disorders (ranging from major depressive disorder to anxiety disorders and obsessive compulsive disorder) found reduced fractional anisotropy as a marker for white matter integrity in frontotemporal and frontoparietal white matter tracts compared with healthy controls.33 The largest clusters of reduced fractional anisotropy incorporated several white matter hyperintensities in frontal and temporal locations.

### Table 2: Results, voxel-based lesion–symptom mapping

<table>
<thead>
<tr>
<th>Anatomic region†</th>
<th>Region size in voxels (n)</th>
<th>Tested voxels (n)</th>
<th>Significant voxels (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forceps major</td>
<td>22285</td>
<td>9537</td>
<td>0</td>
</tr>
<tr>
<td>Forceps minor</td>
<td>35840</td>
<td>5063</td>
<td>0</td>
</tr>
<tr>
<td>Anterior thalamic radiation</td>
<td>43203</td>
<td>13661</td>
<td>0</td>
</tr>
<tr>
<td>Corticospinal tract</td>
<td>27767</td>
<td>5975</td>
<td>15</td>
</tr>
<tr>
<td>Cingulum</td>
<td>13829</td>
<td>1309</td>
<td>0</td>
</tr>
<tr>
<td>Parahippocampal white matter</td>
<td>5234</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inferior fronto-occipital fasciculus</td>
<td>49378</td>
<td>24187</td>
<td>0</td>
</tr>
<tr>
<td>Inferior longitudinal fasciculus</td>
<td>37450</td>
<td>9955</td>
<td>0</td>
</tr>
<tr>
<td>Superior longitudinal fasciculus</td>
<td>59703</td>
<td>29336</td>
<td>1</td>
</tr>
<tr>
<td>Superior longitudinal fasciculus, temporal part</td>
<td>22910</td>
<td>12710</td>
<td>1</td>
</tr>
<tr>
<td>Uncinate fasciculus</td>
<td>15662</td>
<td>4371</td>
<td>0</td>
</tr>
</tbody>
</table>

*Tested and significant voxels for each anatomic region, after correction for age, sex, total white matter hyperintensity volume and multiple testing by applying a false discovery rate.
†Johns Hopkins University diffusion tensor imaging white matter atlas.31

### Table 3: Results, region-of-interest analyses

<table>
<thead>
<tr>
<th>Anatomic region†</th>
<th>β</th>
<th>p value</th>
<th>Region size in voxels (n)</th>
<th>Tested voxels (n)</th>
<th>Significant voxels (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total WMH volume</td>
<td>Model 1</td>
<td>−0.03</td>
<td>0.47</td>
<td>(n = 680)</td>
<td></td>
</tr>
<tr>
<td>Forceps major</td>
<td>Model 1</td>
<td>−0.06</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forceps minor</td>
<td>Model 1</td>
<td>0.05</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior thalamic radiation</td>
<td>Model 1</td>
<td>−0.01</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticospinal tract</td>
<td>Model 1</td>
<td>0.03</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulum</td>
<td>Model 1</td>
<td>−0.02</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior fronto-occipital fasciculus</td>
<td>Model 1</td>
<td>−0.05</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior longitudinal fasciculus</td>
<td>Model 1</td>
<td>−0.04</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior longitudinal fasciculus</td>
<td>Model 1</td>
<td>−0.03</td>
<td>0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior longitudinal fasciculus, temporal part</td>
<td>Model 1</td>
<td>−0.03</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncinate fasciculus</td>
<td>Model 1</td>
<td>−0.01</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MCI = mild cognitive impairment; SCD = subjective cognitive decline; WMH = white matter hyperintensity.

*Results are presented as standardized β. This assumption-free region-of-interest analysis served to identify strategic white matter tracts in which white matter hyperintensity volume is correlated with depressive symptoms, independent of total white matter hyperintensity burden. The Geriatric Depression Scale, as measure of depressive symptoms, was standardized into a z-score. We excluded the tract parahippocampal white matter (Johns Hopkins University atlas31) from our analyses because of the limited white matter hyperintensity in this tract. We first entered age, sex, centre and syndrome diagnosis into a linear regression model (Model 1). If regional volumes showed a significant (p < 0.05) association in Model 1, we added the normalized total white matter hyperintensity volume to the model (Model 2). To check if associations between depressive symptoms and anatomic regions differed according to diagnostic group, we included interaction terms (dummy diagnosis × anatomic region) in the model. When we found an interaction between syndrome diagnosis and anatomic region (p > 0.10), we stratified the results for syndrome diagnosis, and the still is displayed for each diagnostic group separately. When no significant interaction was found, the interaction term was removed from the model and the overall β was reported.
†Johns Hopkins University diffusion tensor imaging white matter atlas.31
‡Significant interaction term; subsequently stratified for syndrome diagnosis.
matter tracts, including the left forceps minor, the anterior thalamic radiation, the inferior fronto-occipital fasciculus and the uncinate fasciculus. A study in patients without dementia but with small-vessel disease found lower white matter integrity in patients with depressive symptoms, particularly in the prefrontal white matter tracts. In contrast, a previous study in a small group of patients with major depressive disorder found increased white matter integrity in the corticospinal tract compared with controls using tractography clustering methods. Previous research suggests that diffusion tensor imaging underestimates fractional anisotropy in regions where fasciculi cross. Because the corticospinal tract is located in an area with crossing fasciculi (i.e., the superior longitudinal fasciculus), these results measured with diffusion tensor imaging should be interpreted with caution. However, our lesion–symptom mapping analyses in a large cohort of memory-clinic patients (including an adjustment for multiple comparisons using FDR correction) also showed an association between the corticospinal tract and depressive symptoms. Our results provide further evidence for a potential role of the corticospinal tract in depressive symptoms. The corticospinal tract is a descending tract of the central nervous system, starting in the cortex and terminating in the spinal cord, and it is known to be involved in controlling movements of the limbs and trunk. It is possible that our findings for the corticospinal tract are related to psychomotor symptoms in depression. We know that depression comprises many combinations of clinical symptoms. Population-based and clinical (in patients with major depressive disorder) studies have investigated the presence of these depressive “subtypes” and suggest the reflection of specific neurobiological biomarkers in particular brain regions between the subtypes. In the present study, we had access only to the total GDS score. Future research with different measures for depressive symptoms is needed to identify the potential presence of depressive subtypes in a memory-clinic population. Consistent with our previous study, we did not find an association between white matter hyperintensity and depressive symptoms in patients with dementia. However, our previous results of a higher propensity for depressive symptoms in patients with subjective cognitive decline and white matter hyperintensity is consistent with the present study, because we found an association between regional white matter hyperintensity in the forceps minor and depressive symptoms in patients with subjective cognitive decline. The forceps minor is a commissural fibre that connects the medial and lateral surfaces of both frontal lobes. It has previously been associated with executive dysfunction and reduced psychomotor speed in patients with vascular brain injury, core cognitive deficits in patients with late-life depression and vascular cognitive impairment. Studies in patients with subjective cognitive decline found subthreshold symptoms of depression and anxiety. Most patients with subjective cognitive decline do not necessarily meet the diagnostic criteria for a psychiatric condition such as major depressive disorder. Affective symptoms in subjective cognitive decline show increased risk of progression to mild cognitive impairment and dementia, suggesting the subthreshold symptoms of depression as a possible manifestation of preclinical Alzheimer disease in these people. Conversely, our analyses in a subgroup with cerebrospinal fluid biomarkers showed that the association between white matter hyperintensities and depressive symptoms is not influenced by Alzheimer disease pathology. To investigate whether factors other than white matter hyperintensity and Alzheimer disease pathology could explain these results, research in other cohorts is needed to provide more evidence. More complex multivariate models (e.g., Bayesian network analysis or multivariate lesion–symptom mapping) might also be of value.

Limitations

Among the limitations of this study is that we used the GDS as a measure of depressive symptoms. Cognitive issues in mild cognitive impairment and dementia may affect the diagnostic accuracy of the GDS. However, the design of the GDS with questions structured in a yes/no format makes it easy to use, even for patients with cognitive impairment. The level and severity of depressive symptoms in this study was relatively low, particularly in patients with dementia, but were consistent with previous studies in memory-clinic populations. Still, this may have reduced the effect sizes and sensitivity to detect associations, despite the large sample size. Second, a relatively high number of patients with subjective cognitive decline used antidepressant medication (20%) compared with patients with mild cognitive impairment (11%) or dementia (12%). The antidepressant medication may have decreased the severity of depressive symptoms and led to lower scores on the GDS, potentially leading to an underestimation of the association between white matter hyperintensity location and depressive symptoms. However, the use of antidepressant medication will be more common in those with higher scores on the GDS, but additional analyses with adjustment for antidepressant medication showed similar results. Finally, the inclusion of our patients at tertiary referral centres and the exclusion of patients with cortical infarcts could limit the generalizability of our findings. On the other hand, the TRACE-VCI cohort is a large memory-clinic cohort of patients with a wide spectrum of vascular brain injury and different levels of cognitive impairment not limited to specific clinical diagnoses such as vascular dementia or Alzheimer disease. In addition, the use of data from different MRI scanners could have influenced the quality of the white matter hyperintensity segmentations and subsequent analyses. However, we have assessed the performance of our segmentation method and it showed no systematic errors across MRI scanners. Nevertheless, additional adjustment for field strength and vendor did not change our results. The use of different MRI scanners could also be seen as a strong point of our study, because it highlights the robustness of our approach and increases the generalizability of our results. Moreover, the large lesion coverage, particularly in the frontoparietal regions, allowed us to include a large number of white matter tracts, leading to greater accuracy and statistical power. We also performed 2 independent hypothesis-free statistical analyses (VLSM and ROI-based linear regression models).
Conclusion

The present study provides further insight into the relationship between white matter hyperintensity location and depressive symptoms by performing a large-scale lesion–symptom mapping study in depressive symptoms.

We have shown that the impact of white matter hyperintensity on depressive symptoms is modest, but appears to be dependent on the location of white matter hyperintensities, particularly in patients with subjective cognitive decline. Our results suggest different etiologies of depressive symptoms in a memory-clinic population with vascular brain injury. Changes in white matter tracts might underlie the occurrence of depressive symptoms in this population.

Acknowledgements: The TRACE-VCI study is supported by Vidi grant 13384 and a gift from ZonMW, Organisation for Health Research and Development, and grant 2010T073 from the Dutch Heart Association to G. Biessels. Research of the VUMc Alzheimer Centre is part of the neurodegeneration research program of Amsterdam Neuroscience. The VUMc Alzheimer Centre is supported by Alzheimer Nederland and Stichting VUMc Fonds. The clinical database structure was developed with funding from Stichting Dioraphte. A. Leeuwis and A. Hooghiemstra are appointed on a grant from The Netherlands CardioVascular Research Initiative: the Dutch Heart Foundation (CVON 2012-06 Heart Brain Connection). F. Barkhof is supported by the NIH Biomedical Research Centre at University College London Hospital.

Affiliations: From the Alzheimer Centre Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands (Leeuwis, Hooghiemstra); Prins, Scheltens, van der Flier); the Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, the Netherlands (Weaver, Biesbroek, Exalto, Biessels); the Image Sciences Institute, University Medical Centre Utrecht, Utrecht, the Netherlands (Kuijf); the Department of Medical Humanities, Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands (Hooghiemstra); the Institutes of Neuroscience and Healthcare Engineering, UCL, London, United Kingdom (Barkhof); the Department of Radiology and Nuclear Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands (Barkhof); and the Department of Epidemiology & Biostatistics, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands (Flier).

Competing interests: N. Prins serves on the advisory board of Boehringer Ingelheim and Probiodrug, and on Abbvie’s DSMB M15-566 trial; has provided consultancy services for Sanofi, Takeda and Kyowa Kirin Pharmaceutical Development; receives research support from Alzheimer Nederland (project number WE.03-2012-02); and is the CEO and co-owner of Brain Research Centre, Amsterdam, the Netherlands. P. Scheltens has acquired grant support (for the institution) from GE Healthcare and Piramal; and in the past 2 years he has received consultancy/speaker fees (paid to the institution) from Medavante, Novartis, Probiodrug, Biogen, Roche, Toyama and EIP Pharma. F. Barkhof serves as a consultant for Biogen-Idec, Janssen Alzheimer Immunotherapy, Bayer-Schering, Merck-Serono, Roche, Novartis, Genzyme and Sanofi-aventis; has received sponsoring from EU-H2020, NWO, SMSR, TEVA, Novartis, Toshiba and Ipi; and serves on the editorial boards of Radiology, Brain, Neuroradiology, MSJ and Neurology. Research programs of W.M. van der Flier have been funded by ZonMW, NWO, EU-FFP, Alzheimer Nederland, Cardiovasculair Onderzoek Nederland, Stichting Dioraphte, Gieskes-Strijbis fonds, Pasman Stichting, Boehringer Ingelheim, Piramal Imaging, Roche BV, Janssen Stellen, Biogen and Combidrugs; all funding is paid to her institution. G. Biessels has been funded by the Dutch Heart Association, ZonMW, the Netherlands Organisation for Health Research and Development and European Union Horizon 2020. No other competing interests declared.

Contributors: A. Leeuwis, N. Weaver, H. Kuijf, W. van der Flier and G. Biessels designed the study. A. Leeuwis, N. Weaver, L. Exalto, H. Kuijf, N. Prins, P. Scheltens, F. Barkhof and G. Biessels acquired the data, which A. Leeuwis, N. Weaver, J.M. Biesbroek, A. Hooghiemstra, N. Prins, P. Scheltens, F. Barkhof, W. van der Flier and G. Biessels analyzed. A. Leeuwis and N. Weaver wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

Members of the TRACE-VCI study group: M.R. Benedictus, J. Bremer, W.M. van der Flier, A.E. Leeuwis, J. Leijenaar, N.D. Prins, P. Scheltens, B.M. Tijms (Alzheimer Centre and Department of Neurology, Amsterdam UMC, location VU University Medical Centre, Amsterdam, The Netherlands); F. Barkhof, M.P. Wattjes (Department of Radiology and Nuclear Medicine, Amsterdam UMC, location VU University Medical Centre, Amsterdam, The Netherlands); C.E. Teunissen (Department of Clinical Chemistry, Amsterdam UMC, location VU University Medical Centre, Amsterdam, The Netherlands); T. Koene (Department of Medical Psychology, Amsterdam UMC, location VU University Medical Centre, Amsterdam, The Netherlands); E. van den Berg, H. van den Brink, G.J. Biessels, J.M.F. Boomsma, L.G. Exalto, D.A. Ferro, C.J.M. Frijs, O. Groeneveld, R. Heinen, S.M. Hering, L.J. Kappelle, Y.D. Reijmers, J. Verver, N.A. Weaver (Department of Neurology, University Medical Centre Utrecht, Utrecht, The Netherlands); J. de Bresser, H.J. Kuijf (Department of Radiology/Imaging Sciences Institute, University Medical Centre Utrecht, Utrecht, The Netherlands); H.L. Koek (Department of Geriatrics, University Medical Centre Utrecht, Utrecht, The Netherlands); J.M.F. Boomsma, H.M. Boss, H.C. Weinstein (Department of Neurology, Onze Lieve Vrouwe Gasthuis (OLVG) West, Amsterdam, The Netherlands); and F. Barkhof (Department of Radiology, National Institute for Health Research (NIHR) and University College London Hospitals NHS Foundation Trust (UCLH) Biomedical research center, London, United Kingdom).

References

Altered white matter connectivity in young people exposed to childhood abuse: a tract-based spatial statistics (TBSS) and tractography study

Lena Lim, PhD*; Heledd Hart, PhD*; Henrietta Howells, PhD; Mitul A. Mehta, PhD; Andrew Simmons, PhD; Kah Mirza, MBBS; Katya Rubia, PhD

Introduction

Brain development is a complex process that is regulated by genes and sculpted by environmental experiences. Although experiential influences affect brain structure and function throughout the lifespan, early childhood experience is particularly crucial, as early stress and exposure to traumatic events have been shown to adversely affect the nature and trajectory of normal brain development.1

Childhood maltreatment, which includes physical, sexual and emotional abuse and neglect, is common in the United Kingdom, with pediatric prevalence of 7%–10%.2 It has been associated with a host of adverse consequences, such as low IQ, abnormal error processing,3 and impaired attention, inhibition, emotion and reward processing.4,5 Large-scale epidemiological studies found that childhood maltreatment was significantly associated with onset of various psychiatric disorders, such as depression and posttraumatic stress disorder (PTSD).6 The psychopathological outcomes associated with childhood maltreatment may be mediated by the disruption of neural underpinnings.7

Background: Childhood abuse is associated with structural brain abnormalities. Few studies have investigated white matter tract abnormalities in medication-naive, drug-free individuals who experienced childhood abuse. We examined the association between childhood abuse and abnormalities in white matter tracts in that population, controlling for psychiatric comorbidities. Methods: We collected diffusion tensor imaging data for age- and sex-matched youth with childhood abuse, psychiatric controls (matched for psychiatric diagnoses) and healthy controls. Tract-specific analysis was conducted using tractography. Tract-based spatial statistics (TBSS) was used to assess group differences in fractional anisotropy (FA) at the whole-brain level. Results: We included 20 youth who experienced childhood abuse, 18 psychiatric controls and 25 healthy controls in our analysis. Tractography analysis showed abuse-specific reduced tract volume in the inferior longitudinal fasciculus (ILF) and inferior frontal-occipital fasciculus (IFoF) in the abuse group relative to both healthy and psychiatric controls. Furthermore, abnormalities in the left IFoF were associated with greater abuse severity. The TBSS analysis showed significantly reduced FA in a left-hemispheric cluster comprising the ILF, IFoF and corpus callosum splenium in the abuse group relative to healthy and psychiatric controls. Limitations: It is unclear to what extent pubertal development, malnutrition and prenatal drug exposure may have influenced the findings. Conclusion: Childhood abuse is associated with altered structure of neural pathways connecting the frontal, temporal and occipital cortices that are known to mediate affect and cognitive control. The abuse-specific deficits in the ILF and IFoF suggest that fibre tracts presumably involved in conveying and processing the adverse abusive experience are specifically compromised in this population.
and in the left motor and somatosensory cortices that mediate sensory functions.22

Compared with the extensive research on grey-matter volume abnormalities in childhood maltreatment, fewer studies have examined white-matter tracts in this population. Brain regions do not function independently; they are interconnected through a complex system of short- and long-range white-matter tracts.15 White matter connectivity regulates the speed and timing of activation across neural networks, which are essential for optimal performance of higher-order tasks that rely on integrated information processing.16

Diffusion tensor imaging (DTI) measures the restricted diffusion of water molecules and provides a more detailed assessment of fibre tracts than conventional MRI and has emerged as a powerful technique for examining structural connectivity.17 Fractional anisotropy (FA), a DTI-derived metric, describes the directionality of water diffusion and may reflect aspects of membrane integrity and myelin thickness, where decreased FA is usually associated with white-matter disruption.18 Tractography facilitates the reconstruction of 3-dimensional trajectories of specific white-matter tracts and probes their microstructure, which allows a more detailed analysis of specific subpopulations of fibres and indirect volumetric indices (e.g., number of streamlines and tract volume).19 These volumetric indices can be indicative of the speed of communication between different brain regions. Tract-based spatial statistics (TBSS), on the other hand, is a fully automated approach that permits a whole-brain analysis of white matter in a voxel-wise manner, which allows the identification of white-matter differences in specific regions beyond a priori-defined tracts.20 Therefore, we used these complementary methods to examine atypical white-matter tracts in youth exposed to childhood abuse.

Stress can affect white-matter tract development, as corticosteroids can suppress the final mitosis of glial cells necessary for myelination.21 Moreover, given the protracted postnatal development timeline of white matter,22 it may be particularly vulnerable to the neurotoxic impact of childhood trauma, especially during certain sensitive periods. Several DTI studies reported that childhood maltreatment was associated with reduced FA in various large white-matter tracts, particularly the inferior fronto–occipital fasciculus (IFoF), which is a direct pathway connecting the occipital, posterior temporal and the OFC areas,23–27 the inferior longitudinal fasciculus (ILF) connecting the occipital with the anterior temporal cortex,23,26,27 which is considered to be an indirect pathway connecting similar brain areas as the IFOF and anteriorly joins the uncinate fasciculus to relay information to the OFC; the superior longitudinal fasciculus (SLF) connecting the Broca and Wernicke areas;24,25,27 the corpus callosum (splenium) connecting the (posterior) left and right cerebral hemispheres;24,25 and the uncinate fasciculus connecting the anterior temporal lobe with the medial and lateral OFC.28

Given that childhood maltreatment is associated with development of psychiatric complications,29 it is crucial to control for these in order to disentangle the effects of maltreatment from psychiatric comorbidities.30 So far, only 3 DTI studies included a psychiatric group without childhood maltreatment;25,30,31 however, those studies used adult samples and focused only on depression, which limits the generalizability of their findings to other psychiatric comorbidities. Furthermore, a number of DTI studies have not measured and/or controlled for drug abuse23,28 and medication use,23,26–28,30 which are known to affect brain structure.32

The aim of the present study was to examine the association between childhood abuse and white-matter tract abnormalities by conducting tract-specific and whole-brain analyses in medication-naive, drug-free youth with documented childhood physical abuse compared with healthy controls. To assess the specificity of the association with abuse, we included a third group of psychiatric controls that was matched with the abuse group on psychiatric comorbidities. Sexual abuse was excluded because it has different effects on brain structure33 and different behavioural and psychiatric consequences.34 It has also been argued that childhood sexual abuse is associated with experiences unique to sexual victimization relative to other abuse experiences; for example, traumatic sexualization, betrayal, stigmatization as well as feelings of guilt and shame may affect victims of sexual abuse differently than victims of other abuse experiences.35 For these reasons, and in order to obtain a more homogeneous group, we included youth exposed only to childhood physical abuse. Nevertheless, it is unrealistic to separate physical abuse from typically co-occurring emotional abuse and neglect because psychological maltreatment would be present in almost all cases of physical maltreatment.36 Hence, it is unlikely that the abused victim would experience severe physical abuse without experiencing at least moderate levels of emotional abuse and neglect concurrently; however, physical abuse does not always co-occur with sexual abuse.

Given that childhood maltreatment is associated with grey-matter volume deficits in OFC–limbic–temporal and occipital visual regions,8,10,13,14 along with abnormalities in the white-matter tracts connecting these regions,23–27 we hypothesized that the abuse group would have white-matter tract abnormalities, particularly of the IFoF and ILF, relative to both the healthy and psychiatric control groups. We also investigated atypical FA in regions beyond our a priori-defined tracts with a whole-brain TBSS analysis.

Methods

Participants

Youth with childhood abuse, psychiatric controls and healthy controls who were right-handed, medication-naive, drug-free and matched for age and sex were assessed by a child psychiatrist (K.M.) using the Development and Well-Being Assessment (DAWBA),37 designed to generate ICD-10 and DSM-IV psychiatric diagnoses. The Strengths and Difficulties Questionnaires (SDQ)38 and Beck Depression Inventory (BDI)39 were also used to provide symptom scores on psychopathology. We assessed IQ using the Wechsler Abbreviated Scale of Intelligence (WASI).40 The Childhood Trauma Questionnaire (CTQ)41 was used to measure the severity of childhood physical, sexual and emotional abuse as well as physical and emotional neglect. Socioeconomic status (SES) was measured by...
2 nonsensitive items (on housing tenure and room occupancy) from the Family Affluence Scale (FAS).42

The 23 youth who experienced physical abuse before the age of 12 years were first recruited through social services and psychiatric clinics. They or their guardians were first asked to provide signed permission to contact social services for written confirmation of official records of physical abuse. The Childhood Experience of Care and Abuse (CECA) interview43 was used to corroborate the CTQ and provide additional information including the age at onset and duration of abuse. Participants scored 13 or higher (i.e., the cut-off for severe/extreme physical abuse)44 on the CTQ physical abuse subscale, and information from the CECA interview and the CTQ were consistent with the official records. Common psychiatric comorbidities included PTSD, depression, anxiety and conduct disorder.

The 20 psychiatric patients who were matched with the abuse group on psychiatric comorbidities but who had no history of childhood maltreatment (scoring below the cut-offs for the respective CTQ subscales)44 were recruited through psychiatric clinics and social services. Patients with PTSD experienced non-abuse-related trauma (e.g., witnessed a murder, experienced a car accident or the death of a loved one).

Participants in the childhood abuse and psychiatric control groups who were recruited from social services did not have any psychiatric diagnoses beforehand, and their family physicians were subsequently notified by the child psychiatrist (K.M.). Those who were recruited from clinics were new clinic patients and had not yet started any treatment, and the diagnoses made using the DAWBA were consistent with the patients’ diagnoses in the clinics. None of the participants was receiving any medication at the time of recruitment and scanning.

The 27 healthy controls with no history of psychiatric illness and childhood maltreatment (scoring below the same cut-offs for the respective CTQ subscales) were recruited through advertisements in the same geographic areas of South London to ensure similar socioeconomic background.

Exclusion criteria for all participants were childhood sexual abuse, drug abuse, learning disability, neurologic abnormalities, epilepsy, IQ below 70 and MRI contraindications. Urine screening for recent drug use was conducted with 10-panel urine drug test integrated cups (T-Cup; Testfield). Participants were also asked about drug use in the previous 4 weeks; most did not use any drugs in the last 4 weeks before the scan and there were no significant group differences (Appendix 1, Table S1, available at jpn.ca/170241-a1). All participants, or their guardians if they were younger than 18 years, provided written informed consent to participate in the study. The study was approved by the local NHS Research Ethics Committee.

Image acquisition and processing

The DTI acquisition procedures are described in Appendix 1. Diffusion data were preprocessed using ExploreDTI (www.exploredti.org).

We assessed group differences in head motion, as this may affect quantitative diffusion measurements. We quantified head motion as the mean volume × volume translation and rotation. This was calculated as the average across the translation or rotation component of the affine registration performed between each volume and the first volume, and t-tests were then performed between the 2 groups for each of the 2 motion measures. As there were no significant group differences in mean translation (\( F_{1,20} = 0.8, p = 0.45 \)) or rotation (\( F_{1,20} = 2.2, p = 0.1 \)), we did not use motion as a nuisance regressor in our results.

Outlier profiles of each diffusion scan were generated using ExploreDTI during the quality check stage of preprocessing, with no difference between groups observed (\( F_{1,20} = 1.20, p > 0.05 \)). All scans were then corrected for head motion using ExploreDTI.

Tractography

We performed virtual dissections of the left and right ILF and IFOF according to previous studies20 (Fig. 1). Regions of interest (ROIs) were delineated in the FA maps of each participant in native space using previously described anatomic guidelines to constrain the whole-brain tractogram.19 Two ROI approaches were used for each tract to show the full extent of white-matter streamlines running through each ROI. Specifically, the ILF was dissected to show streamlines running between the occipital lobe (1 ROI in the coronal plane within the white matter of the occipital lobe) and the temporal pole (1 ROI in the coronal plane within the white matter of the anterior temporal lobe). The IFOF was dissected using the same occipital lobe ROI as used for the ILF and a second ROI delineated in the coronal plane within the external capsule.

Group differences were examined for each measurement (i.e., streamline count, tract volume, FA, mean diffusivity and radial diffusivity) using analysis of variance (ANOVA) with SPSS software (SPSS, Inc.), controlling for IQ, age and sex. Comparisons for specific tracts were considered to be statistically significant if they survived Bonferroni correction for multiple comparisons (\( p < 0.0125, 2 \) tracts for each hemisphere).

Tract-based spatial statistics

Each participant’s FA map was transformed into standard stereotactic space (using the FMRIB58 template), and a mean FA map for the whole sample was used to create the average core “skeleton.” Skeleton images of each participant’s FA map were then produced and projected onto the mean skeleton using a general linear model to identify voxels where FA value differed significantly among these skeletons.20 The design matrix used IQ, age and sex as covariates. Five thousand permutations were applied. The statistical threshold was set at \( p < 0.05 \), fully corrected for multiple comparison using threshold-free cluster enhancement (TFCE) across all white-matter tracts in the whole-brain analysis.

Exploratory correlational analysis

Finally, Pearson correlations were used to explore possible associations between tract-specific measurements and SDQ within each group and with abuse measures (severity, age at onset and duration of abuse) within the abuse group.
Results

Participants

We included 63 youth in our analyses. Of the 23 youth recruited to the abuse group, 3 were excluded owing to MRI motion artifacts, leaving a final sample of 20 participants in that group. Of the 20 recruited psychiatric controls, 2 were excluded owing to motion artifacts, leaving a final sample of 18 patients in that group. Of the 27 healthy controls recruited, 2 were excluded due to motion artifacts, leaving a final sample of 25 participants in that group. The demographic and clinical characteristics of participants are shown in Table 1.

The groups did not differ significantly in age, sex, race or SES, but they differed in IQ, which was expected as this is typical for these populations44 (Table 1). Participants in the childhood abuse group did not mention any head trauma injuries or loss of consciousness from the abuse in the CECA interview. All MRIs were also reviewed by a radiologist, and no traumatic brain injury or incidental findings were discovered. Hence, mild traumatic brain injury is unlikely to affect the findings. Although we selected participants with severe childhood physical abuse, they also experienced marked/severe emotional abuse and neglect (Table 1), which typically co-occur with physical abuse; hence, they seem to adequately represent the childhood abuse population.36

The healthy controls scored significantly lower than the abuse group on the BDI ($p < 0.01$) and all SDQ difficulties subscales ($p < 0.01$), and they scored lower than psychiatric controls on the BDI ($p < 0.001$) and all SDQ difficulties subscales ($p < 0.05$) except for SDQ conduct problems. The abuse group scored significantly higher than psychiatric controls, who did not differ from healthy controls, on the SDQ conduct problems subscale ($p < 0.01$; Table 1).

**Fig. 1:** (A) Tractography reconstructions of the inferior longitudinal fasciculus (ILF) and inferior fronto-occipital fasciculus (IFoF) tracts. (B) Differences in the tract volume of the ILF and IFoF between the childhood abuse group, psychiatric controls and healthy controls. Statistically significant differences between the childhood abuse group and psychiatric and healthy control groups within each tract are indicated with asterisks (*$p < 0.05$; **$p < 0.01$).
Tractography analysis

The abuse group had significantly lower tract volume of the left ILF, right ILF and left IFOF than both healthy ($p < 0.01$) and psychiatric controls ($p < 0.01$) (Table 2, Fig. 1); lower streamline count of the right ILF and left IFOF than both healthy ($p < 0.01$) and psychiatric controls ($p < 0.01$); and lower FA of the left IFOF than healthy controls ($p = 0.01$) (Table 2). There were no significant differences between the healthy and psychiatric controls.

Tract-based spatial statistics analysis

The abuse group, relative to healthy controls, had significantly reduced FA in a left-hemispheric posterior region comprising the ILF, IFOF, splenium of the corpus callosum and the SLF ($F_{1,36} = 16.4, p < 0.001$), which suggests that compromised microstructure of this region may be abuse-specific. The psychiatric controls had marginally lower FA than healthy controls in this region ($F_{1,36} = 3.89, p = 0.06$). There were no significant regions with increased FA for the abuse versus healthy and psychiatric groups.

Exploratory correlational analysis

Reduced FA of the left IFOF was significantly associated with higher CTQ physical neglect ($r = -0.52, p < 0.05$), emotional neglect ($r = -0.48, p < 0.05$) and CTQ total score ($r = -0.50, p < 0.05$) within the abuse group (Appendix 1, Fig. S1). For the healthy controls, FA of the lower left IFOF was significantly associated with higher SDQ emotion ($r = -0.61, p < 0.05$) and

<table>
<thead>
<tr>
<th>Table 1: Demographic characteristics of 20 youth exposed to childhood abuse, 18 psychiatric controls and 25 healthy controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group; mean ± SD or no. (%)</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Age, yr†</td>
</tr>
<tr>
<td>Socioeconomic status</td>
</tr>
<tr>
<td>IQ</td>
</tr>
<tr>
<td>Strengths and Difficulties Questionnaire</td>
</tr>
<tr>
<td>Emotional problems</td>
</tr>
<tr>
<td>Conduct problems</td>
</tr>
<tr>
<td>Hyperactivity</td>
</tr>
<tr>
<td>Peer problems</td>
</tr>
<tr>
<td>Prosocial</td>
</tr>
<tr>
<td>Total difficulties score</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
</tr>
<tr>
<td>Childhood Trauma Questionnaire</td>
</tr>
<tr>
<td>Physical abuse</td>
</tr>
<tr>
<td>Emotional abuse</td>
</tr>
<tr>
<td>Sexual abuse</td>
</tr>
<tr>
<td>Physical neglect</td>
</tr>
<tr>
<td>Emotional neglect</td>
</tr>
<tr>
<td>Age at onset of (physical) abuse, yr</td>
</tr>
<tr>
<td>Duration of (physical) abuse, yr</td>
</tr>
<tr>
<td>Male sex</td>
</tr>
<tr>
<td>DSM-IV Psychiatric diagnosis</td>
</tr>
<tr>
<td>PTSD</td>
</tr>
<tr>
<td>Depression</td>
</tr>
<tr>
<td>Anxiety disorders</td>
</tr>
<tr>
<td>Social phobia</td>
</tr>
<tr>
<td>Panic disorder</td>
</tr>
<tr>
<td>ADHD</td>
</tr>
<tr>
<td>ODD/CD/other disruptive behaviours</td>
</tr>
</tbody>
</table>

ADHD = attention-deficit/hyperactivity disorder; CA = childhood abuse group; CD = conduct disorder; HC = healthy controls; ODD = oppositional defiant disorder; PC = psychiatric control group; PTSD = post-traumatic stress disorder.

*Bonferroni-corrected.
†Age range 13–20 years.

---

J Psychiatry Neurosci 2019;44(4)
Table 2: Measurements of the inferior longitudinal fasciculus and inferior fronto-occipital fasciculus tracts

<table>
<thead>
<tr>
<th>Tract</th>
<th>Childhood abuse n = 20</th>
<th>Psychiatric controls n = 18</th>
<th>Healthy controls n = 25</th>
<th>CA v. HC comparisons</th>
<th>CA v. PC comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ILF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamlines</td>
<td>426 270</td>
<td>614 364</td>
<td>726 396</td>
<td>3.23 0.047</td>
<td>4.47 0.041 CA &lt; HC</td>
</tr>
<tr>
<td>Tract volume</td>
<td>1465 544</td>
<td>1953 673</td>
<td>2149 591</td>
<td>6.23 0.004*</td>
<td>9.02 0.005 CA &lt; HC</td>
</tr>
<tr>
<td>FA</td>
<td>0.495 0.021</td>
<td>0.499 0.022</td>
<td>0.501 0.021</td>
<td>0.42 0.66</td>
<td>0.45 0.51</td>
</tr>
<tr>
<td>MD</td>
<td>0.793 0.026</td>
<td>0.790 0.033</td>
<td>0.786 0.021</td>
<td>1.35 0.27</td>
<td>3.66 0.07</td>
</tr>
<tr>
<td>RD</td>
<td>0.553 0.027</td>
<td>0.547 0.031</td>
<td>0.543 0.025</td>
<td>0.96 0.39</td>
<td>2.29 0.14</td>
</tr>
<tr>
<td>Right ILF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamlines</td>
<td>339 282</td>
<td>666 393</td>
<td>719 404</td>
<td>7.15 0.002*</td>
<td>11.9 0.001 CA &lt; HC</td>
</tr>
<tr>
<td>Tract volume</td>
<td>1180 684</td>
<td>2091 695</td>
<td>2132 650</td>
<td>12.2 &lt; 0.001*</td>
<td>16.7 &lt; 0.001 CA &lt; HC</td>
</tr>
<tr>
<td>FA</td>
<td>0.480 0.040</td>
<td>0.486 0.023</td>
<td>0.493 0.022</td>
<td>0.76 0.47</td>
<td>1.13 0.29</td>
</tr>
<tr>
<td>MD</td>
<td>0.792 0.027</td>
<td>0.786 0.032</td>
<td>0.780 0.020</td>
<td>0.67 0.52</td>
<td>1.54 0.22</td>
</tr>
<tr>
<td>RD</td>
<td>0.561 0.037</td>
<td>0.553 0.033</td>
<td>0.545 0.024</td>
<td>1.14 0.33</td>
<td>2.16 0.15</td>
</tr>
<tr>
<td>Left IFoF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamlines</td>
<td>406 330</td>
<td>849 436</td>
<td>960 509</td>
<td>10.0 &lt; 0.001*</td>
<td>12.6 0.001 CA &lt; HC</td>
</tr>
<tr>
<td>Tract volume</td>
<td>1762 872</td>
<td>2776 599</td>
<td>2860 677</td>
<td>14.3 &lt; 0.001*</td>
<td>14.8 &lt; 0.001 CA &lt; HC</td>
</tr>
<tr>
<td>FA</td>
<td>0.499 0.027</td>
<td>0.510 0.026</td>
<td>0.516 0.024</td>
<td>3.20 0.048</td>
<td>7.41 0.010 CA &lt; HC</td>
</tr>
<tr>
<td>MD</td>
<td>0.810 0.029</td>
<td>0.794 0.032</td>
<td>0.796 0.020</td>
<td>3.05 0.05</td>
<td>5.93 0.020 CA &gt; HC</td>
</tr>
<tr>
<td>RD</td>
<td>0.561 0.036</td>
<td>0.542 0.036</td>
<td>0.540 0.027</td>
<td>3.74 0.03</td>
<td>8.46 0.006 CA &gt; HC</td>
</tr>
<tr>
<td>Right IFoF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamlines</td>
<td>409 332</td>
<td>676 384</td>
<td>706 356</td>
<td>2.74 0.07</td>
<td>2.14 0.15</td>
</tr>
<tr>
<td>Tract volume</td>
<td>1694 831</td>
<td>2352 830</td>
<td>2390 733</td>
<td>3.61 0.03</td>
<td>2.91 0.10</td>
</tr>
<tr>
<td>FA</td>
<td>0.496 0.017</td>
<td>0.502 0.022</td>
<td>0.509 0.022</td>
<td>1.59 0.21</td>
<td>3.57 0.07</td>
</tr>
<tr>
<td>MD</td>
<td>0.801 0.030</td>
<td>0.802 0.037</td>
<td>0.793 0.020</td>
<td>0.20 0.82</td>
<td>0.47 0.50</td>
</tr>
<tr>
<td>RD</td>
<td>0.558 0.026</td>
<td>0.554 0.035</td>
<td>0.543 0.023</td>
<td>0.74 0.48</td>
<td>2.25 0.14</td>
</tr>
</tbody>
</table>

CA = childhood abuse group; FA = fractional anisotropy; HC = healthy controls; ILF = inferior longitudinal fasciculus; IFoF = inferior fronto-occipital fasciculus; MD = mean diffusivity; PC = psychiatric control group; RD = radial diffusivity; SD = standard deviation.

*Indicates values that survive Bonferroni correction for multiple comparisons.

To our knowledge, this is the first DTI study to examine the association between documented childhood abuse and alterations in the structure of neural pathways in medication-naive, drug-free youth, controlling for psychiatric comorbidities by the inclusion of a psychiatric control group. This is crucial to elucidate the effects of abuse independently from effects associated with psychiatric comorbidities or medication and drug abuse. As hypothesized, the abuse group had significantly reduced white-matter tract volume in the bilateral ILF and left IFoF compared with both healthy and psychiatric controls. At the whole-brain level, the abuse group also had significantly reduced FA in a left-hemispheric posterior region comprising the ILF, IFoF, splenium of the corpus callosum and SLF relative to both healthy and psychiatric controls. Reduced FA of the left IFoF, which was also found in the tractography results, correlated with greater abuse severity in the abuse group. This suggests differences exist not only at the microstructural level as measured by FA, but also at the volumetric level of the entire tract. Thus, differences in the white-matter level of the entire tract were also observed.

Discussion

To our knowledge, this is the first DTI study to examine the association between documented childhood abuse and alterations in the structure of neural pathways in medication-naive, drug-free youth, controlling for psychiatric comorbidities by the inclusion of a psychiatric control group. This is crucial to elucidate the effects of abuse independently from effects associated with psychiatric comorbidities or medication and drug abuse. As hypothesized, the abuse group had significantly reduced white-matter tract volume in the bilateral ILF and left IFoF compared with both healthy and psychiatric controls. At the whole-brain level, the abuse group also had significantly reduced FA in a left-hemispheric posterior region comprising the ILF, IFoF, splenium of the corpus callosum and SLF relative to both healthy and psychiatric controls. Reduced FA of the left IFoF, which was also found in the tractography results, correlated with greater abuse severity in the abuse group. This suggests differences exist not only at the microstructural level as measured by FA, but also at the volumetric level of the entire tract. Thus, differences in the white-matter level of the entire tract were also observed.
matter of the ILF and IFoF, particularly in the left hemisphere, was specifically related to the abuse experience. Moreover, reduced FA of the left IFoF was significantly associated with higher SDQ emotion and peer problems in the healthy controls, reinforcing the association between the IFoF and emotional and social behaviours.

The ILF is a ventral associative bundle that mediates the fast transfer of visual signals from the visual areas to the amygdala and hippocampus, and neuromodulatory back-projections from the amygdala to early visual areas, enhancing the visual processing of emotionally significant stimuli. It is a key component of the visual–limbic pathway involved in facial affect recognition and visual perception. The finding of an abuse-specific reduced white-matter microstructure of the ILF extends the findings of earlier studies that found decreased FA of the ILF in adolescents exposed to early neglect and in young adults with childhood maltreatment, where the decreased FA was furthermore related to poorer visual learning and memory in neglected adolescents and with longer duration of abuse.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>MNI coordinates</th>
<th>Cluster size</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left inferior longitudinal fasciculus/inferior fronto-occipital fasciculus/splenium of the corpus callosum/superior longitudinal fasciculus</td>
<td>–31, –69, –1</td>
<td>678</td>
<td>0.02</td>
</tr>
</tbody>
</table>

MNI = Montreal Neurological Institute; TFCE = threshold-free cluster enhancement
The right hemisphere is particularly dominant for negative emotional processing in most individuals.40 Thus, it seems that abuse exposure affects corticolimbic regions involved in emotional regulation and specifically targets the visual–limbic pathway involved in the emotional processing of (aversive) visual information. Given that the abuse experience has both visual and auditory components, the left ILF may also have been compromised as it is involved in language processing.49 Interestingly, studies suggest that fearful facial expressions alone activate the right amygdala, while fearful facial expressions combined with fearful voices activate the left medial temporal gyrus.39 Hence, the combined exposure to fearful faces and voices during a typical severe abuse episode may have disrupted the normal development of both the left and right ILF.

The IFoF, which overlaps spatially and functionally with the ILF, connects the ventral occipital, posterior temporo-basal areas to the frontal lobe (inferior frontal, dorsolateral prefrontal and emotion-related OFC regions) and runs parallel to the ILF in its occipital course.51 Hence, it is also involved in facial affect recognition,45 visual and semantic processing, and in multimodal sensorimotor integration.52 Altered microstructure of the IFoF is also consistent with the findings of earlier studies that reported lower FA of the IFoF in adolescents exposed to early neglect53 and in individuals with childhood maltreatment.24,25 The association between abuse experience and microstructure of the IFoF is further underpinned by the present findings of significant negative correlation between abuse severity and FA of the left IFoF.

The splenium of the corpus callosum interconnects the left and right occipital and inferior temporal cortices.54 These regions form the ventral visual stream with reciprocal connections with the hippocampus and emotion-related structures such as the amygdala and OFC.55 The splenium has a protracted myelination trajectory from birth to early adulthood with an accelerated growth during middle childhood that accompanies the development of visual–spatial integration.54 It is involved in the integration of somatosensory and emotional visual information in the 2 hemispheres.55 Our findings also support earlier studies that found reduced FA of the splenium in individuals exposed to childhood maltreatment.24,25

Childhood maltreatment has been associated with abnormal development of the sensory systems that relay adverse sensory experiences. For instance, studies reported structural deficits in the occipital-lingual regions in children with maltreatment56 and psychosocial deprivation,57 in women who experienced childhood sexual/physical abuse,53 and in young adults who witnessed domestic violence during childhood.53 These findings suggest that the sensory systems that process and interpret adverse sensory inputs may be altered by the abuse experience, reflecting an adaptive response of the developing brain to protect the child from highly hostile environmental conditions by gating sensory experiences and processing related to the abuse.33

Similarly, childhood maltreatment is associated with structural deficits in the emotion-related OFC6–10 and amygdala regions,59 along with functional abnormalities in frontolimbic regions while processing fearful or angry faces.59,60 Therefore, besides impairment in these individual regions, the findings of white-matter alterations in the ILF and IFoF tracts further suggest disruptions in visual–limbic–OFC pathways mediating sensory integration and cognitive or emotion regulation to sensory stimuli, which may also underlie the neurological deficits in emotion and reward processing61,62 observed in childhood maltreatment.

Given that large-scale epidemiological and longitudinal studies have consistently shown that childhood maltreatment is linked developmentally to psychiatric disorders,29 it is crucial to control for these in order to disentangle the effects of maltreatment from psychiatric comorbidities.10 Therefore, the specificity of the present findings of differences in the ILF and IFoF at both the microstructural and volumetric levels relative to a psychiatric control group in particular extends the findings of previous studies and suggests that these neural pathways are specifically compromised in abused individuals.

The human brain is a highly plastic organ that is continually modified by experience and undergoes changes across the lifespan. The individual neural regions and circuits mature at different rates and have different windows of vulnerability to effects of traumatic stress, with increased vulnerability ascribed to a period of rapid maturation.63 Studies suggest that the maturation of neuronal circuits of the human visual cortex may extend beyond infancy into childhood, with significant development in visual spatial integration between 5 and 14 years of age.64 Given that the ILF, IFoF and splenium show rapid development from childhood with FA increase peaking at early adulthood,64 the visual–limbic pathways may be more susceptible to impairment in individuals with early adversities. Thus, our findings of an association between childhood maltreatment and altered structure of these late developing visual–emotional processing tracts suggests an environmentally triggered disturbance in the normal development of these pathways that may underlie the emotional problems that arise as a consequence of early adversities.

Limitations

Among the strengths of this study are that all participants were medication-naïve and drug-free, and their abuse experience was carefully assessed and corroborated by social service records. Also, we included a psychiatric control group to determine the specificity of childhood abuse in our findings. The inclusion of a childhood abuse group without any psychiatric disorders would have provided a more robust means of determining abuse-specific abnormalities; however, such a “pure” group would not be representative of the general childhood abuse populations, as large-scale epidemiological and longitudinal studies have consistently reported that childhood maltreatment is linked developmentally to psychiatric disorders,29 and a meta-analysis further reported a causal relationship between nonsexual childhood maltreatment and a range of mental disorders.65 For the tractography analysis, multiple comparison correction was performed for the number of tracts only and not for the number of diffusion measures, as these are not independent from each other and Bonferroni correction would thus have been too conservative.
It is unclear to what extent pubertal development, malnutrition, prenatal drug exposure and presence of current life stressors may have influenced the findings. The moderate sample size of the present study warrants replication in larger samples of youth in future studies. The SES measure used is limited, as it does not provide information on parents’ income and education; however, youth often have difficulties reporting this information. Although we recruited participants exposed to childhood physical abuse, it is unrealistic to separate physical abuse from typically co-occurring emotional abuse and neglect; hence, many participants in the abuse group also suffered from emotional abuse and neglect.

Conclusion

Using medication-naive, drug-free, carefully assessed age- and sex-matched groups of youth exposed to childhood abuse and psychiatric controls matched on psychiatric comorbidities, we found that childhood abuse is associated with altered microstructure of neural pathways connecting the OFC limbic, temporal and occipital visual regions. The abuse-specific abnormalities of the ILF and IFOF visual–limbic pathways may underlie the abnormal emotional regulation to sensory stimuli in victims of abuse.

Acknowledgments: L. Lim and H. Howells were supported by the Reta Lila Weston Trust for Medical Research and the Kids Company UK. L. Lim was also supported by the Lee Kong Chian School of Medicine, Nanyang Technological University Singapore Fellowship Grant (grant number L0491050). H. Howells was supported by the Wellcome Trust (grant number 103759/Z/14/Z). A. Simmons and K. Rubia received research support from the UK Department of Health via the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) for Mental Health at South London and the Maudsley NHS Foundation Trust and Institute of Psychiatry, Psychology and Neuroscience, King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors thank all the individuals who participated in this study and their families, as well as the staff of Kids Company London for their help with recruitment.

Affiliations: From the Department of Child & Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK (Lim, Hart, Mirza, Rubia); the Lee Kong Chian School of Medicine, Imperial College – Nanyang Technological University Singapore Fellowship Grant (grant number L0491050). H. Howells was supported by the Wellcome Trust (grant number 103759/Z/14/Z). A. Simmons and K. Rubia received research support from the UK Department of Health via the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) for Mental Health at South London and the Maudsley NHS Foundation Trust and Institute of Psychiatry, Psychology and Neuroscience, King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors thank all the individuals who participated in this study and their families, as well as the staff of Kids Company London for their help with recruitment.

Competing interests: M. Mehta has acted as a consultant for Cambridge Cognition and Lundbeck and has received fees from Shire for contribution towards education. K. Mirza has received research and educational grants from GlaxoSmithKline and Shire pharmaceuticals; served on the advisory boards of Janssen, Eli Lilly and Shire pharmaceuticals; and received honoraria for speaking at conferences organized by Janssen, Eli Lilly and Shire pharmaceuticals. K. Rubia has received speaker’s honoraria from Lilly and Shire. No other competing interests declared.

Contributors: L. Lim, H. Hart, M. Mehta, K. Mirza and K. Rubia designed the study. L. Lim, H. Hart and A. Simmons acquired the data, which L. Lim, H. Hart, H. Howells and K. Rubia analyzed. L. Lim and K. Rubia wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

References

2. NSPCC. Child abuse and neglect in the UK today; 2012.


Audi Canada is pleased to present a preferred incentive for Canadian physicians.

Enjoy additional lease/finance and purchase incentives starting from $750 to $4,000.

For more information visit www.AudiProfessional.ca

The all-new 2019 Audi A6

The all-new MMI® touch response features two large touch screens to give you total control.

Audi pre sense® uses front, side, and rear sensors to deliver a full 360° approach to safety.

The Audi virtual cockpit keeps you on track with Google Earth Navigation.