Predicting Alzheimer's dementia from heterogeneous patterns of neurodegeneration and functional connectivity

ANGELA TAM
Integrated Program in Neuroscience
McGill University, Montreal
August 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

© Angela Tam, 2018

Abstract

There is a large field of research dedicated to the development of biomarkers for an early diagnosis of Alzheimer's disease (AD). Predicting AD dementia within an individual, especially at a prodromal stage like mild cognitive impairment (MCI), is complicated by the vast amount of heterogeneity present in populations. This thesis explores heterogeneity in brain organization with magnetic resonance imaging (MRI) in order to develop biomarkers to identify individuals who will progress to AD dementia. Chapter 1 provides a brief introduction to the problem at hand and lists the specific aims of this thesis. Chapter 2 provides a review of the literature of biomarker development for AD, with a focus on MRI-based studies and prediction of cognitive trajectories with machine learning. In chapter 3, we present a study that explored whether there were functional connections in resting-state networks that could consistently discriminate between patients with MCI and cognitively normal older adults in the face of heterogeneity from methodological procedures. We identified functional connections that were robustly altered in the default mode network and the cortical-strial-thalamic loop, albeit with small to medium effect sizes, in MCI patients compared to controls in several independent datasets. We also provide sample size estimates to obtain adequate statistical power in a multisite study setting. Chapter 4 presents a study describing resting-state networks at various spatial resolutions in a heterogeneous sample of older adults with and without cognitive impairment. In chapter 5, we describe a study in which we developed a signature based on brain atrophy patterns and cognitive deficits that is highly predictive of future progression to AD dementia in a subgroup of individuals with MCI. By harnessing the heterogeneity in brain structure, we were able to achieve higher positive predictive values and specificity compared to previous works at predicting progression to dementia from the MCI stage. Lastly, a discussion of the contributions and future developments from these studies is presented in chapter 6. This thesis provides novel insights into the heterogeneity of structural and functional brain organization and the use of MRI as a tool to develop biomarkers for AD.

Résumé

Il y a un grand domaine de recherche dédié à l'identification précoce de la maladie d'Alzheimer (MA). La prédiction de la démence liée à la MA chez un individu, surtout dans ceux qui sont atteints des troubles cognitifs légers (TCL), est compliquée à cause de la variabilité au niveau des populations. Cette thèse explore l'hétérogénéité dans l'organisation du cerveau avec l'imagerie par résonance magnétique (IRM) pour l'objectif de développer des biomarqueurs capables d'identifier des individus qui seront atteints de la démence liée à la MA. Chapitre 1 introduit le problème et les objectifs spécifiques de la thèse. Chapitre 2 consiste en un survol de la littérature du développement des biomarqueurs pour la MA, avec un emphase sur les études d'IRM et la prédiction des trajets cognitifs avec l'apprentissage machine. Chapitre 3 offre un étude où on examine la présence des connections fonctionnels dans les réseaux d'état de repos qui pourraient discriminer systématiquement entre les patients avec des TCL et les personnes âgées cognitivement sains, malgré l'hétérogénéité des procédures méthodologiques. On a identifié des connections fonctionnels dans le réseau mode par défaut et le circuit entre le cortex, striatum et thalamus qui étaient altérés de manière robuste dans les patients avec des TCL comparés aux contrôles dans plusieurs jeux de données indépendants. Nous offrons aussi des estimations des tailles d'échantillon pour obtenir le pouvoir statistique adéquat dans un contexte de recherche multisite. Chapitre 4 présent un étude qui décrit les réseaux d'état de repos à travers de nombreux résolutions spatiaux dans un échantillon hétérogène de personnes âgées avec et sans des troubles cognitifs. Dans chapitre 5, on a développer une signature basée sur les patrons d'atrophie du cerveau et déficits cognitifs qui est hautement prédictive de la démence liée à la MA dans un sous-groupe d'individus avec TCL. En employant l'hétérogénéité de la structure du cerveau, on rapport des valeurs prédictive positives et des spécificités plus hausses que les études antérieures qui ont visé à prédire la progression vers la démence de l'étage des TCL. Finalement, le chapitre 6 porte sur une discussion des contributions et développements futurs de ces études. Cette thèse apporte des nouveaux apprentissages par rapport à l'hétérogénéité de l'organisation structurelle et fonctionnelle du cerveau et l'utilité de l'IRM comme un outil pour le développement des biomarqueurs de la MA.

Contents

List of Figures

List of Tables

List of Abbreviations

Aβ	**A**myloid β
AD	**A**lzheimer's **D**isease
ADNI	**A**lzheimer's **D**isease **N**euroimaging **I**nitiative
aMCI	**a**mnestic **M**ild **C**ognitive **I**mpairment
APOE	**APO**lipoprotein **E**
APP	**A**myloid **P**recursor **P**rotein
BASC	**B**ootstrap **A**nalysis of **S**table **C**lusters
CN	**C**ognitively **N**ormal
DMN	**D**efault **M**ode **N**etwork
fMRI	**f**unctional **M**agnetic **R**esonance **I**maging
HPS	**H**ighly **P**redictive **S**ignature
LDA	**L**inear **D**iscriminant **A**nalysis
MCI	**M**ild **C**ognitive **I**mpairment
MRI	**M**agnetic **R**esonance **I**maging
MSTEPS	**M**ultiscale **STEP**wise **S**election
PET	**P**ositron **E**mission **T**omography
pMCI	**p**rogressive **M**ild **C**ognitive **I**mpairment
PSEN1	**P**re**SEN**ilin **1**
PSEN2	**P**re**SEN**ilin **2**
sMCI	**s**table **M**ild **C**ognitive **I**mpairment
SVM	**S**upport **V**ector **M**achine
VBM	**V**oxel-**B**ased **M**orphometry

Acknowledgements

I would first like to thank my supervisor Dr. Pierre Bellec for his support. I am without a doubt a significantly more competent scientist than when I first started my Ph.D. and I could not have done that without his investment in me. I cannot have imagined a better mentor and I will always be grateful for the lessons and opportunities that Pierre provided. I would like to thank my co-supervisor Dr. John Breitner for his support, guidance, and advice throughout my degree. I thank my SIMEXP labmates (past and present), Dr. Aman Badhwar, Sebastian Urchs, Yassine Benhajali, Perrine Ferré, Clara Moreau, Amal Boukdhir, Jonathan Armoza, Dr. Pierre-Olivier Quirion, Phil Dickinson, Dr. Yu Zhang, for the many laughs and for creating a fun and stimulating work environment. I would especially like to thank Dr. Christian Dansereau and Dr. Pierre Orban for their mentorship, and I was lucky to have them as my role models. I'm grateful for the entire SIMEXP team and I'm going to miss everyone terribly. I also thank Dr. Marie-Élyse Lafaille-Magnan for her solidarity and discussions. I also give thanks to my friends outside of academia for their tremendous support and encouragement. I want to thank Jacob Vogel for the long days and nights we spent working alongside each other and everything else in between. Lastly, I thank my family, especially my parents, my brother Brandon, and my grandparents, for their love and the many home-cooked meals that helped me get through every challenge, scientific or otherwise, as a better person.

To my parents. . . I'm sorry I got a
job in Singapore so I probably
won't come back for graduation.

"Reality is not what you perceive; it's what the methods and tools of science reveal."

Neil deGrasse Tyson

Contribution of authors

As lead author of the studies described in chapters 3 through 5, I took the lead in executing the experiments. This included study design, data preprocessing, quality assessment, statistical analysis, visualization and interpretation of results, and writing of the manuscripts. I also received much appreciated and valuable help from my co-authors, outlined below.

Chapters 3 and 4

- Christian Dansereau: data preprocessing and quality control
- AmanPreet Badhwar: writing of the manuscript
- Pierre Orban: visualization of results
- Sylvie Belleville: review of the manuscript
- Howard Chertkow: advice for methods
- Alain Dagher: data collection
- Alexandru Hanganu: data collection
- Oury Monchi: data collection
- Pedro Rosa-Neto: data collection and curation
- Amir Shmuel: data collection
- Seqian Wang: data curation
- John Breitner: writing of the manuscript
- Pierre Bellec: study design, interpretation of results, and writing of the manuscript

Chapter 5

- Christian Dansereau: design of methods
- Yasser Iturria-Medina: design of methods
- Sebastian Urchs: design of methods
- Pierre Orban: design of methods
- John Breitner: review of manuscript
- Pierre Bellec: study design, interpretation of results, and writing of manuscript

Chapter 1

Introduction

1.1 General context

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of dementia. As AD is an age-related disorder, more and more individuals will develop AD as the population ages. For instance, there are currently 53 million adults aged 65 and older in the United States in 2018. This segment of the American population is projected to increase to 88 million by 2050, followed by an expected doubling of the number of AD cases (Alzheimer's Association, 2018). Under current circumstances for diagnosis and treatment, the cost of medical and long-term care expenditures for American individuals who will develop AD in 2018 is projected to be $47.1 trillion USD (Alzheimer's Association, 2018). AD clearly presents a public health crisis that requires imminent solutions.

Unfortunately, there is currently no disease-modifying drug for AD that can reverse or slow down the course of the disease. The failures for clinical trials of disease-modifying agents have been blamed on lack of efficacy, among other reasons (Cummings, Morstorf, and Zhong, 2014). The lack of efficacy could be due to drugs being tested too late in the disease process when irreversible neurodegeneration has already occurred in clinical trial participants (Aisen, Vellas, and Hampel, 2013). The pathology of AD develops over decades prior to the emergence of clinical symptoms. Individuals who develop AD dementia typically experience an intermediate prodromal phase called mild cognitive impairment (MCI), where some cognitive deficits are apparent but not severe enough to impede daily function. There is therefore a push to target pre-symptomatic individuals with disease-modifying agents to try to prevent disease progression. A second reason for a lack of demonstrated efficacy could be heterogeneity within clinical populations. AD exists in several forms, and an individual's risk for developing AD depends on a variety of factors including, but

not limited to, genetics, cardiovascular health, and brain reserve. This heterogeneity may explain why inclusion criteria for clinical trials have had low to moderate positive predictive value for diagnosing MCI subjects who would develop AD dementia (Visser, Scheltens, and Verhey, 2005). The failure of clinical trials at the MCI stage in order to prevent dementia may be partially attributed to the incorrect inclusion of individuals who will not develop AD dementia. The development of a biomarker for early and precise diagnosis that would account for this heterogeneity could greatly improve patient selection for clinical trials and could identify high-risk individuals for earlier interventions.

1.2 Objectives

A promising tool that is already widely used in clinical settings to aid in the diagnosis of AD is magnetic resonance imaging (MRI). The overall objective of this thesis was to explore heterogeneity within MRI-based biomarkers in order to discover a brain signature that is highly predictive of Alzheimer's dementia. Such a signature could include, for example, atrophy in the medial temporal lobes and specific nodes of the default mode network, dysconnectivity within those regions, and cognitive deficits. We focused on two MRI modalities, structural and functional MRI, and examined their potential to explain cognitive outcomes. Specific objectives from each chapter that presents original research are described below.

Chapter 3 objectives

Chapter 3 of this thesis assessed the robustness of resting-state connectivity derived from functional MRI as an early biomarker for AD, in the face of heterogeneity from different image acquisition and diagnostic protocols. To this end, we combined multiple independent datasets of resting-state functional magnetic resonance images from MCI and cognitively normal subjects to test the consistency of functional connectivity differences between these two diagnostic groups.

Chapter 4 objectives

Chapter 4 was a companion paper to Chapter 3 and described resting-state network parcellations that are present in a large heterogeneous population of older adults with or without MCI.

Chapter 5 objectives

Chapter 5 of this thesis characterized the variability in brain atrophy patterns, derived from structural MRI, and characterized the heterogeneity to develop a signature that has high positive predictive value at predicting incipient AD dementia in MCI subjects. We achieved this by identifying subtypes of brain atrophy in AD and control subjects. We then applied a machine learning algorithm to classify AD and controls based on brain atrophy subtypes and cognitive test scores. This resulted in a signature that was common to AD patients but not represented in controls. We then transfered the predictive model to identify a subset of individuals who carried the signature and who would progress to AD dementia from a dataset of MCI subjects.

Chapter 2

Review of the literature

2.1 Alzheimer's disease

Alzheimer's disease (AD) is an age-related neurodegenerative disorder and the leading cause of dementia, representing between 60% to 80% of cases (Alzheimer's Association, 2018). Dementia is a syndrome characterized by deficits in cognition, such as memory or language problems, that impair an individual's ability to live independently and to perform daily activities. Dementia represents an end-stage of AD as cognitive impairments are only apparent after the pathology of AD has developed within an individual for decades. The pathological markers of AD are amyloid beta (Aβ) plaques and neurofibrillary tau tangles in the brain. The "amyloid cascade hypothesis" (Hardy and Selkoe, 2002) is currently the most widely accepted disease model for AD. This hypothesis posits that excess Aβ in the brain is the primary driving force behind AD. It is believed that an imbalance between producing and clearing Aβ leads to an increase in the Aβ peptide and plaque formation, which disrupts synaptic function and accelerates tau phosphorylation and neurofibrillary tangle formation. This eventually causes synaptic failure and neuronal death, which subsequently results in dementia. Note that this hypothesis is yet to be proven and that it might not fully capture the complexity of AD. The cascade of events leading to AD dementia is illustrated by Figure 2.1.

Amyloid precursor protein (APP), presenilin 1 (PSEN1), and presenilin 2 (PSEN2) have been identified as causative genes associated with a familial form of AD that presents with an early onset of symptoms (before the age of 60 to 65 years). AD cases arising from mutations in APP, PSEN1, and PSEN2 account for about 1% of all AD cases (Bekris et al., 2010). Mutations in these genes all lead to an overproduction of Aβ42 (Bekris et al., 2010), a form of Aβ that is more prone to aggregating together

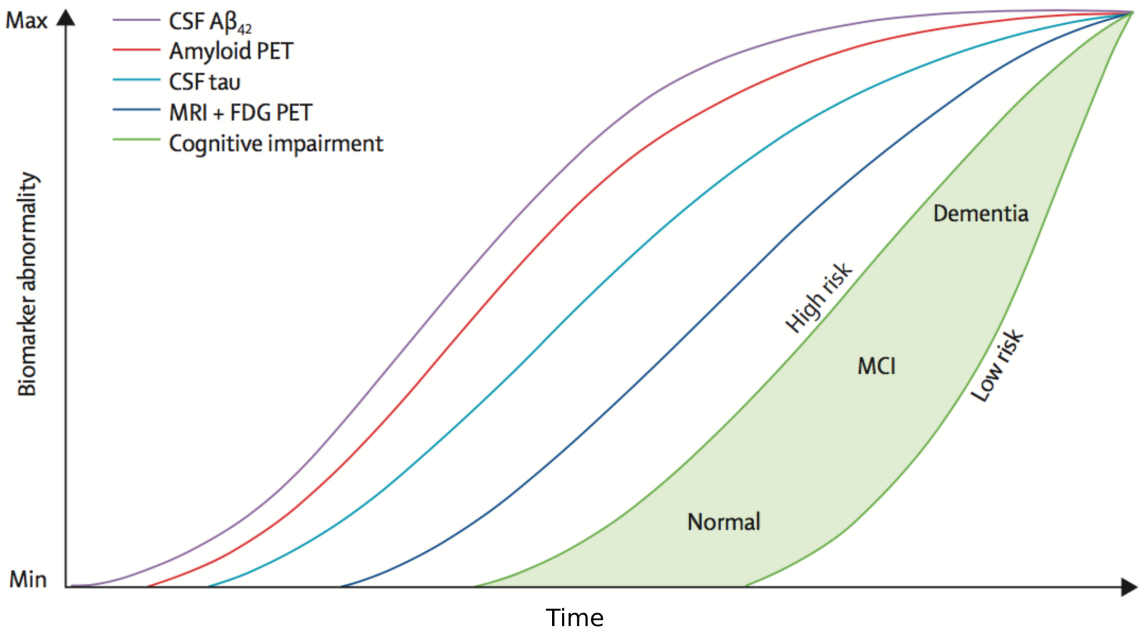

Figure 2.1: Model of dynamic biomarkers of the Alzheimer's disease pathological cascade. Neurodegeneration is measured by FDG PET and structural MRI, which are drawn concordantly (dark blue). By definition, all curves converge at the top right-hand corner of the plot, the point of maximum abnormality. Cognitive impairment is illustrated as a zone (light green-filled area) with low-risk and high-risk borders. People who are at high risk of cognitive impairment due to Alzheimer's disease pathophysiology are shown with a cognitive impairment curve that is shifted to the left. By contrast, the cognitive impairment curve is shifted to the right in people with a protective genetic profile, high cognitive reserve, and the absence of comorbid pathological changes in the brain, showing that two patients with the same biomarker profile at a given time can have different cognitive outcomes. Aβ=amyloid β. CSF=cerebrospinal fluid. FDG=fluorodeoxyglucose. MCI=mild cognitive impairment. MRI=magnetic resonance imaging. PET=positron emission tomography. Figure and adapted caption from Jack et al. (2013).

and forming the amyloid plaques that are associated with AD (Snyder et al., 1994). These cases of AD are called autosomal dominant forms of AD.

The large majority of AD cases present with a later onset of symptoms (after the age of 60 to 65 years). Unlike the autosomal dominant forms of AD, the cause for late-onset AD is less clear. The greatest risk factors for this form of AD are age, carrying the $\epsilon 4$ allele of the apolipoprotein (APOE) gene, and having a family history of AD. Of these risk factors, age is the greatest risk factor as the prevalence of AD increases significantly with advancing age: 3% of individuals between 65-75 years have AD, but this number jumps to 32% in those aged 85 years or older (Hebert et al., 2013). So far, a polymorphism in the APOE gene is the single greatest genetic risk factor for developing late-onset AD. The APOE gene has three common alleles: APOE $\epsilon 2$, APOE $\epsilon 3$, and APOE $\epsilon 4$. APOE $\epsilon 3$ is the most common allele (Farrer et al., 1997). Compared to $\epsilon 3$, carrying the $\epsilon 2$ allele may confer a protective effect against developing AD, while the $\epsilon 4$ allele is a major risk factor (Farrer et al., 1997). The risk of AD increases by a factor of 2.84 for each $\epsilon 4$ allele that an individual carries, which means that an individual with a 4/4 genotype is more than 8 times as likely to develop AD compared to subjects with 2/3 or 3/3 genotypes (Corder et al., 1993). However, there are still other genetic (and non-genetic) factors that remain to be discovered as having first-degree relatives afflicted with AD dementia also increases one's risk for AD, even after accounting for APOE genotype (Martinez et al., 1998). This thesis will focus on the late-onset form of AD.

2.1.1 Mild cognitive impairment

Mild cognitive impairment (MCI) is a clinical entity that encompasses the grey area between intact cognitive function and clinical dementia (Petersen et al., 2014). Individuals with MCI experience a degree of cognitive impairment that is not so severe as to impede activities of daily living. MCI was previously seen as a transient stage on the path toward dementia. However, the heterogeneity within this clinical entity actually makes prognosis rather challenging. A meta-analysis reported that only up to 36% of MCI patients will develop AD dementia within a span of two years (Ward et al., 2013). This shows that while MCI can be a significant risk factor for AD dementia, many MCI patients will remain cognitively stable.

Different subtypes of MCI have been described based on symptomatic expression. Often, MCI patients are categorized as amnestic MCI, patients who have memory impairment, non-amnestic MCI, those who have impairments in non-memory domains, and multi-domain MCI, patients with concurrent impairments in multiple kinds of cognition (Petersen, 2003). Evaluating an individual's prognosis can be improved by placing the patient in a specific subgroup because certain subgroups may have

higher or lower risks for dementia. Amnestic MCI subjects are more likely to be diagnosed with dementia than non-amnestic MCI (Aggarwal et al., 2005; Busse et al., 2006), but the group at highest risk of conversion is multi-domain MCI (Tabert et al., 2006). Of multi-domain MCI subjects, 59% progressed after two years of follow-up, while only 18% of amnestic single domain MCI progressed, and 70% of non-amnestic single domain MCI actually improved. This is evidence that outcomes can vary substantially among MCI subpopulations.

2.1.2 Diagnosis of Alzheimer's disease

It used to be argued that the only definitive method for determining whether an individual has AD is through post-mortem neuropathological examinations. During such an assessment, a pathologist will look for amyloid plaques and neurofibrillary tangles inside the brain of an individual to determine whether the cause for one's dementia is AD. In living individuals, only clinical diagnoses of "probable AD dementia" or "possible AD dementia" are possible. A clinical diagnosis of probable or possible AD dementia is based on cognitive assessments and a patient's history (McKhann et al., 2011). Brain imaging with magnetic resonance imaging (MRI) is often used for differential diagnosis from other conditions that could cause dementia.

Interestingly, clinical diagnoses often do not match with pathology, which reveal that the previous gold standard of a pathological diagnosis may be an oversimplification. About a third of cognitively normal older adults had significant AD pathology when examined after death (Bennett et al., 2006). Beach et al. (2012) found that 17% of clinically diagnosed AD dementia patients did not meet criteria for a neuropathological definition of AD. The same study also found 33% of cases were unclassifiable because neuropathological guidelines were too restrictive (Beach et al., 2012). It is becoming increasingly clear that AD is a complex and heterogeneous disorder, and we need to study the variability within populations if we are to ever reach the level of personalized medicine.

In fact, it has been shown that there are neuropathologically defined subtypes of Alzheimer's disease. Autopsy studies have characterized AD subtypes based on the density and distribution of neurofibrillary tangles in the cortex and hippocampus (Murray et al., 2011). Murray et al. (2011) described "typical", "hippocampal sparing", and "limbic predominant" cases of AD. Compared to typical cases, the hippocampal sparing subtype had less neurofibrillary tangles in the hippocampus than in the cortex, whereas the limbic predominant subtype had more hippocampal tangles than in the cortex (Murray et al., 2011). Given that tau causes neurodegeneration (Spillantini and Goedert, 2013), the variability in tau deposition may account for the heterogeneity seen in brain atrophy and symptomatic expression in AD.

2.2 Biomarkers of Alzheimer's disease

A biomarker is an objective, accurate, and reproducible measure of indications of a medical state that can be observed from outside a patient (Strimbu and Tavel, 2010). According to the National Institutes of Health Biomarkers Definitions Working Group, a biomarker can be used to evaluate normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (Biomarkers Definitions Working Group et al., 2001). There are several biomarkers that have been or are currently being developed to detect or track the progression of AD. These biomarkers include measures of molecular pathology, neuroimaging, and cognitive assessments.

2.2.1 Amyloid and tau markers

Biomarkers for Aβ and tau are regarded as core biomarkers for AD as these two proteins form AD pathology. Aβ and tau can be quantified through the cerebrospinal fluid (CSF). In AD, there is a marked reduction of Aβ in the CSF, which reflects greater Aβ plaque formation in the brain (and by consequence, less diffusion of Aβ into the CSF) (Blennow et al., 2010). Tau, on the other hand, is present in higher concentrations in the CSF in AD, and reflects the degree of neuronal degeneration (Blennow et al., 2010). More recently, positron emission tomography (PET) imaging is increasingly being used to quantify Aβ and tau. While these biomarkers directly measure the underlying pathology of AD, their usage outside of research settings for diagnosis is currently not recommended (McKhann et al., 2011).

2.2.2 Magnetic resonance imaging markers

Magnetic resonance imaging (MRI), on the other hand, is a widely used clinical tool. With different MRI techniques, we can measure brain structure and function. This thesis will focus on two forms of MRI: T1-weighted structural imaging and T2*-weighted task-free functional MRI (fMRI).

2.2.2.1 Structural imaging of brain atrophy

Structural imaging of the brain through T1-weighted MRI has become an important aspect in the clinical assessment of AD as it can capture neurodegeneration and because the degree of atrophy measured with MRI correlates well with tau deposition and cognitive deficits (Frisoni et al., 2010). Brain atrophy can be measured with different metrics extracted from structural MRI. Two common measures are grey matter volume and cortical thickness. It is well known that extensive neuronal

loss occurs in the medial temporal lobes in AD patients (Jack et al., 1997). Medial temporal lobe atrophy may be an early marker of AD as it can also be seen in MCI patients (McDonald et al., 2009). As the disease progresses and clinical symptoms become more severe, atrophy typically spreads to the parietal and frontal lobes, while leaving the sensorimotor and visual cortices relatively intact (McDonald et al., 2009); see Figure 2.2. Although this represents a stereotypical pattern of AD-related neurodegeneration, there is actually substantial variation in atrophy patterns among individuals, which is likely due to heterogeneity in tau pathology. A number of works have reproduced a subtype characterized by primarily medial temporal lobe atrophy, a second subtype characterized by parietal, occipital, and lateral temporal lobe atrophy, and a third subtype with diffuse atrophy (Noh et al., 2014; Hwang et al., 2016; Park et al., 2017; Varol et al., 2017); see Figure 2.3. These atrophy subtypes correspond well to the neuropathological subtypes that were characterized by Murray et al. (2011).

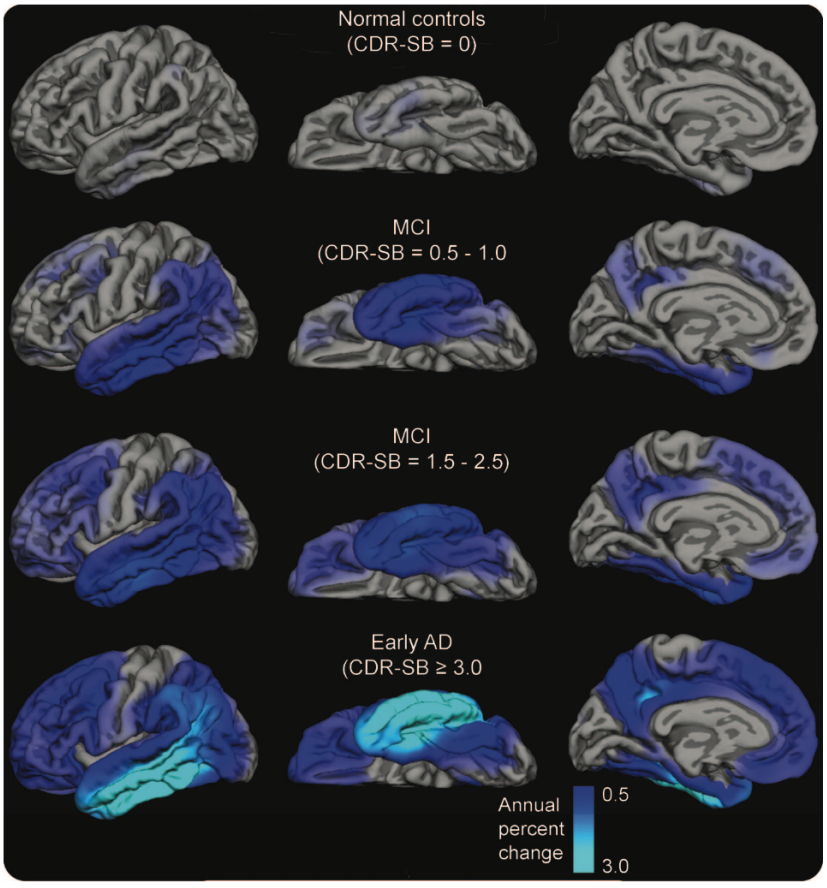

Figure 2.2: Annual atrophy rates as a function of degree of clinical impairment (i.e., baseline Clinical Dementia Rating Sum of Boxes score [CDR-SB]). Mean atrophy rates are represented as a percent change in neocortical volume and mapped onto the lateral (left), ventral (middle), and medial (right) pial surface of the left hemisphere. These data demonstrate that atrophy rates are most prominent in posterior brain regions early in the course of disease, spreading to anterior regions as the level of impairment increases, with relative sparing of sensorimotor regions. MCI=mild cognitive impairment; AD=Alzheimer disease. Figure and caption from McDonald et al. (2009).

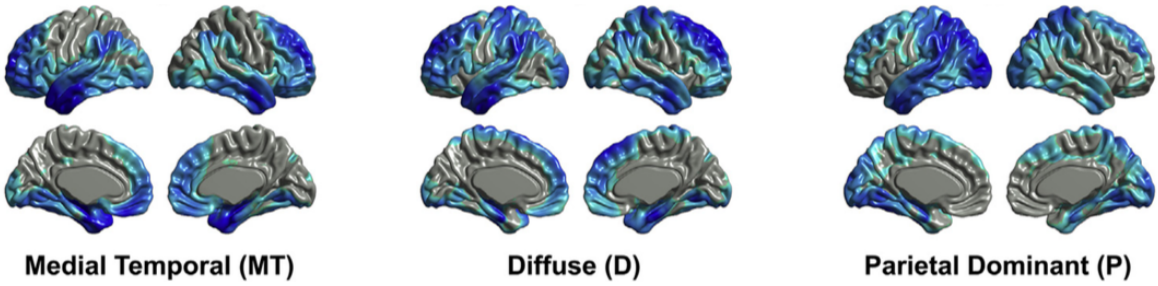

Figure 2.3: Images of cortical thinning patterns in three subgroups of Alzheimer's disease dementia patients, compared with cognitive normal subjects. Gray areas indicate brain regions showing no statistical significant in cortical thickness compared to control groups. Darker blue areas indicate areas with more cortical thinning in AD patients compared to controls. Figure and adapted caption from Hwang et al. (2016).

Similarly to the different types of MCI, subtypes of atrophy may carry varying degrees of risk of developing dementia. The subgroup with parietal, occipital, and lateral temporal lobe atrophy tended to be younger and had the worst cognition (Noh et al., 2014; Park et al., 2017), while the medial temporal lobe atrophy subtype contained a greater proportion of carriers of the APOE4 allele (Varol et al., 2017), and the diffuse atrophy subtype had the least Aβ and tau burden (Varol et al., 2017). Expressing specific patterns of brain atrophy may be predictive of cognitive trajectories. Zhang et al. (2016), for example, found that subjects with more temporal lobe atrophy had steeper declines in memory, while subjects with diffuse atrophy across the cortex had steeper declines in executive function. Additionally, Dong et al. (2017) described subjects with widespread brain atrophy, and especially those who also express greater loss in the temporal lobes, progress faster from MCI to AD dementia compared to subjects with more localized atrophy (Dong et al., 2017).

2.2.2.2 Resting-state functional neuroimaging

Functional magnetic resonance imaging (fMRI) is a tool that measures oxygenated blood flow, which we can use as a proxy for neural activity. A localized increase in neural activity is directly reflected in a localized increase in the fMRI signal (Logothetis et al., 2001), as oxygenated blood will rush towards areas with sustained neural activity to replenish depleting resources of the active neurons. Resting-state connectivity in fMRI captures the coherence of spontaneous fluctuations in blood oxygenation across different brain regions. This can enable us to examine brain networks, formed from regions that are said to be functionally connected when they exhibit the same fluctuations in blood flow. Common resting-state networks are shown in Figure 2.4.

In research contexts, MRI images typically go through some preprocessing before any statistical analyses are done. For example, T1-weighted structural images will usually undergo as a spatial normalization so that the brains of all subjects within a

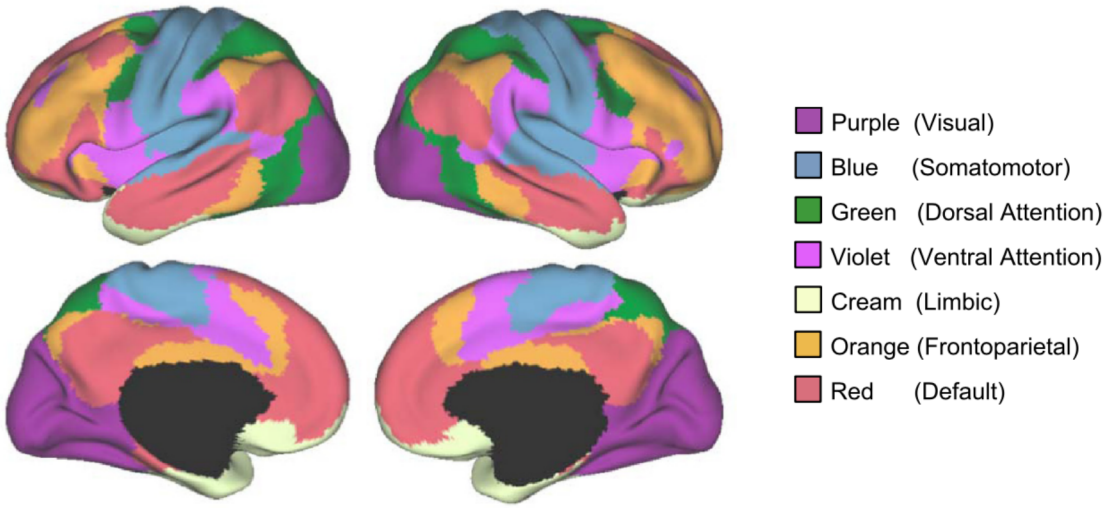

Figure 2.4: A 7-network cortical parcellation from resting-state fMRI. Common names associated with each network in the neuroimaging literature are included in parentheses. Figure and caption adapted from Yeo et al., 2011.

dataset will be in the same space. This enables us to make inferences a group level about localized effects, such as atrophy in a specific brain region. In resting-state fMRI, not only does spatial normalization occurs to align the brains of different subjects as in T1-weighted image preprocessing, but other preprocessing steps are applied to correct for potential artifacts that could influence subsequent analyses. Physiological activity (e.g. heart beat, respiration, head motion) may impact the resulting fMRI signal, for example. Common preprocessing steps for resting-state fMRI images include: 1) slice timing correction to correct for sampling different layers of the brain at slightly different times, 2) rigid-body motion estimation to correct for motion, 3) coregistration of the fMRI image to a reference space, 4) resampling the fMRI image to the reference space, 5) regression of nuisance covariates from the fMRI time series, such as slow time drifts, white matter signal, and ventricle signal, and 6) spatial smoothing to improve the signal to noise ratio.

The seminal work of Greicius et al. (2004) found that functional connectivity in the default mode network (DMN) (illustrated as the "red" network in Figure 2.4) may be able to distinguish AD patients from controls, which spurred interest in developing resting-state fMRI as a biomarker tool for AD. Since then, many groups have reported decreased functional connectivity between nodes of the DMN in AD or MCI patients compared to controls (Bai et al., 2009; Koch et al., 2012; Liang et al., 2012; Sorg et al., 2007; Zhang et al., 2010) (also see Badhwar et al. (2017) for a meta-analysis of 37 studies). The value of resting-state fMRI as an AD biomarker may be even greater when we consider the evidence that Aβ deposition occurs in DMN regions (Buckner et al., 2005) and that AD-related neurodegeneration preferentially spreads

through the DMN (Seeley et al., 2009). There is also evidence that aberrant DMN connectivity appears prior to any measurable amyloid plaques in cognitively normal APOE ϵ4-carrying individuals (Jones et al., 2016; Sheline et al., 2010b), which may place DMN function at the epicentre of the AD cascade.

Although there is a semblance of consistency in the resting-state fMRI literature in AD, substantial variation exists. For example, when we examine studies individually, we can find reports of increased connectivity between the middle temporal lobe and other DMN areas in MCI (Qi et al., 2010), as well as reports of decreased connectivity between these same regions (Bai et al., 2009), and others have reported no significant differences between MCI and controls (Koch et al., 2012). While low sample sizes can be a simple explanation for this variability, there are also methodological differences that should be considered. For example, variability from subject selection or image acquisition may make comparisons of results across independent studies difficult. The robustness of functional connectivity changes in AD populations must be first tested if resting-state fMRI is to be used as a potential biomarker.

2.3 Predicting clinical trajectories with machine learning

There is a growing body of literature dedicated to using machine learning to predict clinical outcomes. Broadly speaking, machine learning can be divided into different subcategories: supervised and unsupervised. The learning is considered to be supervised when input-output labels are given to the predictive model during the training phase (Jain, Murty, and Flynn, 1999) so that the resulting classifier will be able to label new unseen data. A classification problem that would use a supervised algorithm could be the automatic diagnosis of AD based on features such as clinical symptoms and brain atrophy. Commonly used supervised machine learning algorithms include support vector machines, logistic regression, decision trees, and k-nearest neighbours (Dreiseitl and Ohno-Machado, 2002).

Unsupervised learning, on the other hand, is completely data-driven and aims to find meaningful structure in data that is unlabeled or uncategorized (Jain, Murty, and Flynn, 1999). Data mining, for example, finding subgroups within data when it is unknown how the data is divided, would use an unsupervised algorithm. Clustering algorithms (e.g. k-means, fuzzy k-means, hierarchical, mixture Gaussian) are popular examples of unsupervised learning (Jain, Murty, and Flynn, 1999).

Ensemble methods in machine learning combine multiple algorithms with the aim to achieve better predictive performance compared to any of the individual

member algorithms (Dietterich, 2000). An ensemble of classifiers will classify new unseen data by taking a weighted vote of the predictions by the individual classifiers (Dietterich, 2000). Popular ensemble methods include Bayesian averaging, bagging, and boosting (Dietterich, 2000).

One advantage to applying machine learning to the medical field is the potential to produce a tool that could automatically detect progression to AD dementia with high predictive power in a data-driven fashion. Metrics that are commonly used to evaluate the performance of a predictive model include accuracy, positive predictive value, sensitivity and specificity. Accuracy represents the proportion of subjects that were correctly identified as either positives or negatives. Positive predictive value is the proportion of true positives out of all subjects that were classified as positive, and it depends on the prevalence of a condition in a population. Sensitivity is the proportion of individuals with the condition (e.g. progressive MCI) who were correctly identified as positives by the model. Specificity is the proportion of individuals without the condition (e.g. stable MCI) who were correctly identified as negatives by the model.

This thesis will focus on building a predictive model with MRI markers and cognition to classify MCI subjects who will progress to AD dementia (i.e. progressive MCI (pMCI)) and those who will remain cognitively stable (i.e. stable MCI (sMCI)).

2.3.1 Cognition-based predictive models

Given the central role cognitive testing plays in the diagnosis of dementia, a number of studies have used neuropsychological assessments to identify pMCI. Cognitive test scores have tended to be the strongest predictors of incipient dementia when compared to other types of features, such as imaging or CSF (Korolev et al., 2016; Eckerström et al., 2013).

Memory impairment is an important feature in AD dementia and has often been thought of as a defining symptom of the disease. Previous works have shown that certain delayed recall tasks have high predictive power to identify progression to dementia, with the best performances to date reported at $\geq$92% accuracy, $\geq$92% sensitivity, and $\geq$83% specificity (Kluger et al., 1999; Eckerström et al., 2013; Irish et al., 2011). A meta-analysis of 21 studies also found that verbal episodic memory tests consistently report over 70% sensitivity and specificity and are among the features with the highest predictive power at identifying pMCI (Belleville et al., 2017).

While memory tests appear to have good predictive value by themselves, better performance can be achieved by combining multiple kinds of neuropsychological assessments. Belleville et al. (2017) found that the specificity and sensitivity for

individual cognitive domains can vary and recommend that a combination of neuropsychological measures could yield high predictive accuracy with a good balance between specificity and sensitivity. The combination of memory tests with language and/or executive function tests seems particularly promising (Dickerson et al., 2007; Belleville et al., 2017; Belleville et al., 2014). The best performance by a multi-domain model, so far, used logistic regression and reported 87.7% accuracy, 86.2% sensitivity, and 88.9% specificity (Belleville et al., 2014).

2.3.2 Imaging-based predictive models

There has also been a large amount of research with respect to predictive models using features derived from structural MRI. Of these features, there are generally three types that have been used in the literature for AD classification: voxel-segmented tissue probability maps, cortical thickness, and pre-defined regions of interest.

With voxel-segmented tissue probability maps, patterns of atrophy can be measured by examining grey matter, white matter, and/or CSF, which could be quantified by voxel-based morphometry (VBM) (Ashburner and Friston, 2000) for example. Previous works that have used voxel-segmented tissue probability maps as features for classifying AD can be further divided into two categories: 1) studies that used voxels from the entire brain tissue maps as features, and 2) studies that used reduced features extracted from the tissue probability maps. Studies that have used voxels across entire brain with different classifiers (e.g. support vector machine (SVM), regularized logistic regression, linear discriminant analysis (LDA)) have achieved 81-95% accuracy in AD classification and 62-75% accuracy in predicting progression to AD from the MCI stage (see Rathore et al. (2017) for a review). One of the issues with using the whole brain is that the number of features is typically much larger than the number of available data points (i.e. subjects). In this situation, an algorithm will often fail to generalize to new data. Feature-reduction methods (e.g. principle component analysis) can be used to reduce the dimensionality to overcome this issue and generate fewer, yet potentially more meaningful, features. The dimensionality of whole brain maps can also be reduced by parcellating the brain into regions defined by an atlas and use the grey matter volume for each parcel as a feature. Studies that have used reduced features have reported 76-95% accuracy in AD classification and 66-82% accuracy in predicting progressive MCI (Rathore et al., 2017).

While cortical thickness is typically extracted from all vertices on the cortical surface, the majority of classification studies using cortical thickness as features have used feature reduction methods or extracted features from pre-defined atlases. This is likely due to the high dimensionality problem that is present in cortical thickness studies in a similar way to studies that use voxel-segmented tissue probability maps.

Park et al. (2012) reported that using reduced cortical features yielded better classification accuracy than using raw vertices by a margin of 10-13%, which suggests that many features across the entire brain are uninformative and classification is improved when such features are removed. Studies using cortical thickness as features have achieved 87-95% accuracy for AD classification and 71-81% accuracy for predicting progressive MCI (Rathore et al., 2017).

Of studies that make classifications based on pre-defined regions of interest, most have used hippocampal morphometry (e.g. shape, volume) as features. These studies have reported 81-94% accuracy for AD classification and 74-80% accuracy for predicting progressive MCI (Rathore et al., 2017).

2.3.3 Impact of machine learning algorithms on predictive performance

Different machine learning algorithms may perform differently on the same task (e.g. one algorithm may be more accurate than another). The choice in algorithm may therefore account for some of the variability in AD classification, in addition to the variability due to the choice of features.

For the studies that used cortical surface-based AD classification frameworks that were reviewed in Rathore et al. (2017), half used SVM, while a bit less than half used LDA, and one used logistic regression. The studies that used SVM reported, on average, higher classification accuracy than the LDA studies (by a margin of 2-3%) when distinguishing between AD and CN, although the highest accuracy of 95% was reported by Desikan et al. (2009), which used logistic regression. The result from Desikan et al. (2009) outperforms the highest performing SVM by Wee et al. (2013), which reported 92.3% accuracy, by a margin of 3-4%.

Of the studies that used VBM-based AD classification frameworks that were reviewed in Rathore et al. (2017), the majority (90%) used SVM as their classification algorithm. As such, it is difficult to assess if there are significant differences in performance across different algorithms with VBM features because there are few examples of non-SVM models and there was a very wide range in classification accuracy (while predicting AD vs CN) of the SVM models that overlapped with the non-SVM models.

Lastly, of the studies reviewed in Rathore et al. (2017) that used hippocampal volumes as the feature set to classify AD vs CN, 5 out of 6 used SVM, while one used logistic regression. The SVM models reported accuracies ranging from 88-95%, while the logistic regression model reported 81% accuracy, which may suggest that with this smaller feature set, SVM may yield better accuracies.

The most compelling evidence illustrating the impact of machine learning algorithms on AD classification comes from a recent work by Samper-González et al. (2018). Samper-González et al. (2018) directly compared different algorithms (SVM, logistic regression, and random forests) to automatically classify AD patients with T1-weighted MRI and PET images measuring glucose uptake as features across three different large datasets. In this work, the authors found that linear SVM and L2-regularized logistic regression gave similar performances, but they both outperformed random forests, and no influence was found from feature types (Samper-González et al., 2018).

2.3.4 Identifying high-risk subgroups

Given the complexity of predicting incipient AD dementia in individuals with MCI, 80% accuracy may seem good enough. However, relying on accuracy as a performance metric can be misleading. A model could have high accuracy yet moderate positive predictive value if the prevalence for disease is low. Relatively speaking, few MCI subjects progress to dementia (up to 36% within two years (Ward et al., 2013)). In a situation like this where there are imbalanced classes, a model may favour the larger class by ignoring the misclassification of the smaller class in order to obtain a high accuracy. This issue has been largely ignored in research to predict AD. Despite promising accuracy in imaging-based models to predict AD progression, the positive predictive values remain moderate, ranging from 50 to 75% across the literature (Dansereau et al., 2017). This implies that up to half of the individuals who were labeled as progressors in previous studies did not actually progress to AD dementia.

An ideal model to predict conversion to AD dementia would have both high sensitivity and high specificity, so that it would be able to detect all progressor MCI and reject all stable MCI individuals. In practice, there is a trade off between sensitivity and specificity, such that optimizing for one will negatively affect the other. Selecting the optimal balance will depend on the specific problem at hand. In the case of predicting AD progression, the pathophysiological heterogeneity of the disease will prevent highly accurate prediction linking brain features to clinical trajectories. Given this heterogeneity, it is necessary to sacrifice sensitivity to focus on identifying a subgroup of individuals with similar brain abnormalities and cognitive deficits who would be at the highest risk of progressing to AD dementia. This would result in higher specificity and positive predictive value. Characterizing the heterogeneity that is present in the AD spectrum to predict cognitive trajectories in a high-risk subgroup can have important implications for situations that require highly precise predictions, like subject selection for clinical trials or identifying individuals for an early diagnosis.

Chapter 3

Common effects of amnestic mild cognitive impairment on resting-state connectivity across four independent studies

Angela Tam, Christian Dansereau, AmanPreet Badhwar, Pierre Orban, Sylvie Belleville, Howard Chertkow, Alain Dagher, Alexandru Hanganu, Oury Monchi, Pedro Rosa-Neto, Amir Shmuel, Seqian Wang, John Breitner, Pierre Bellec

3.1 Preface

At the onset of this study, there had been many groups that reported resting-state functional connectivity differences between cognitively normal individuals and patients with MCI and AD dementia. Some of these findings, especially those reporting dysfunction in the default mode network in patient groups, seemed consistent across the literature (see Badhwar et al. (2017) in Appendix A for a meta-analysis which I co-authored), but there have been conflicting reports. Inconsistencies in this field may be due to low sample sizes, as the majority of the literature had performed their experiments with 20 subjects per group, leading to a lack of statistical power. However, there is a substantial amount of heterogeneity with respect to methodology (e.g. subject selection, image acquisition and preprocessing, analytical strategies) that may account for a lack of consensus. This study tested for the consistency of functional connectivity alterations in MCI patients compared to controls across a multi-site setting, where independent cohorts were combined together despite variations in scanner manufacturer, imaging protocol, and subject recruitment. The purpose was

to assess whether there are functional connections that are robust enough in the face of heterogeneity that could be used to consistently discriminate between normal aging and MCI. A biomarker must be reliable enough to overcome methodological variation. This work therefore has significant implications for biomarker development. While we did find consistent functional connectivity alterations across the different MCI cohorts when compared to controls, heterogeneity across the samples was substantial. It is apparent from the results of this study that small samples cannot produce findings that will generalize well to the population. This article was published in Frontiers in Aging Neuroscience in 2015[1].

3.2 Abstract

Resting-state functional connectivity is a promising biomarker for Alzheimer's disease. However, previous resting-state functional magnetic resonance imaging studies in Alzheimer's disease and amnestic mild cognitive impairment (aMCI) have shown limited reproducibility as they have had small sample sizes and substantial variation in study protocol. We sought to identify functional brain networks and connections that could consistently discriminate normal aging from aMCI despite variations in scanner manufacturer, imaging protocol, and diagnostic procedure. We therefore combined four datasets collected independently, including 112 healthy controls and 143 patients with aMCI. We systematically tested multiple brain connections for associations with aMCI using a weighted average routinely used in meta-analyses. The largest effects involved the superior medial frontal cortex (including the anterior cingulate), dorsomedial prefrontal cortex, striatum, and middle temporal lobe. Compared with controls, patients with aMCI exhibited significantly decreased connectivity between default mode network nodes and between regions of the cortico-striatal-thalamic loop. Despite the heterogeneity of methods among the four datasets, we identified common aMCI-related connectivity changes with small to medium effect sizes and sample size estimates recommending a minimum of 140 to upwards of 600 total subjects to achieve adequate statistical power in the context of a multisite study with 5-10 scanning sites and about 10 subjects per group and per site. If our findings can be replicated and associated with other established biomarkers of Alzheimer's disease (e.g. amyloid and tau quantification), then these functional connections may be promising candidate biomarkers for Alzheimer's disease.

[1] https://doi.org/10.3389/fnagi.2015.00242

3.3 Introduction

Resting-state connectivity in functional magnetic resonance imaging (fMRI) captures the spatial coherence of spontaneous fluctuations in blood oxygenation. Resting-state fMRI is a promising technique that may be useful as an early biomarker for Alzheimer's disease (AD), a neurodegenerative process that develops over decades before patients suffer from dementia. The possibility that disturbed resting-state connectivity may be an early marker for AD is supported by studies of mild cognitive impairment (MCI), a disorder characterized by objective cognitive deficits without dementia, i.e., without impairment in activities of daily living, and more specifically by studies of amnestic MCI (aMCI), the most common subtype of MCI characterized by memory deficits (Petersen et al., 2001). These studies showed altered functional connectivity in MCI compared with cognitively normal elderly (CN) (Bai et al., 2009; Liang et al., 2012; Sorg et al., 2007; Wu et al., 2014), but they relied on small sample sizes (n 40) and differed in many aspects of their protocols, e.g. recruitment and image acquisition procedures. If resting-state fMRI is to serve as a useful biomarker of AD, or any pathology, for clinical practice or research, we must determine if changes in functional connectivity differences between groups of subjects are robust to such variation in study protocols. Therefore, we sought to identify brain connections that showed consistent MCI-related changes across multiple independent studies. If such connections exist, they may be used as targets to be examined alongside other established AD biomarkers (e.g. amyloid and tau measures) in order to validate resting-state fMRI's potential as a biomarker for AD.

Resting-state connectivity studies have consistently found decreased connectivity between nodes within the default mode network (DMN) in patients with AD or MCI compared with CN (Bai et al., 2009; Koch et al., 2012; Liang et al., 2012; Sorg et al., 2007; Zhang et al., 2010). Less consistent are reports of alterations in the executive attentional, frontoparietal, and anterior temporal networks (Agosta et al., 2012; Gour et al., 2011; Liang et al., 2012; Sorg et al., 2007; Wu et al., 2014; Zhang et al., 2010) due to the literature's bias towards investigating the DMN. Further inconsistencies can be found in some studies that have reported increased connectivity between the middle temporal lobe and other DMN areas in MCI (Qi et al., 2010), while others have reported decreased connectivity between these same regions (Bai et al., 2009) and others have reported no significant differences between MCI and CN (Koch et al., 2012).

One obvious explanation for such inconsistency may be these studies' small sample sizes resulting in low statistical power (Kelly et al., 2012). Beyond this, however, there are other methodological differences that may compromise the comparison of

results across independent studies. For example, the criteria for recruiting subjects with MCI, e.g. Petersen (2004) vs. NIA-AA recommendations (Albert et al., 2011) may differ among studies. Different study samples may also reflect different socio-cultural characteristics of recruiting sites, e.g., ethnicity, language, diet, socioeconomic status. The fMRI measurements themselves can also be affected by differences in details of the image acquisition such as scanner make and model (Friedman, Glover, and Fbirn Consortium, 2006), sequence parameters such as repetition time, flip angle, or acquisition volume (Friedman and Glover, 2006), experimental design such as eyes-open/eyes-closed (Yan et al., 2009) or experiment duration (Van Dijk et al., 2010), and scanning environment such as sound attenuation measures (Elliott, Bowtell, and Morris, 1999), room temperature (Vanhoutte, Verhoye, and Linden, 2006), or head-motion restraint techniques (Edward et al., 2000).

To identify robust changes in resting-state connectivity between aMCI and CN, we implemented a meta-analysis of four independent resting-state fMRI datasets (ADNI2 and three small single-site studies) using a weighted average implemented by Willer, Li, and Abecasis (2010). Rather than relying on a priori target regions or connections, we leveraged the large sample size to perform a systematic search of brain connections affected by aMCI, an approach termed a "connectome-wide association study" (Shehzad et al., 2014). In addition, we relied on functionally-defined brain parcellations using an automated clustering procedure and we explored the impact of the number of brain clusters (called resolution) on observed differences (Bellec et al., 2015).

3.4 Methods

3.4.1 Participants

We combined data from four independent studies: the Alzheimer's Disease Neuroimaging Initiative 2 (ADNI2) sample, two samples from the Centre de recherche de l'institut universitaire de gériatrie de Montréal (CRIUGMa and CRIUGMb), and a sample from the Montreal Neurological Institute (MNI) (Wu et al., 2014). All participants gave their written informed consent to engage in these studies, which were approved by the research ethics board of the respective institutions, and included consent for data sharing with collaborators as well as secondary analysis. Ethical approval was also obtained at the site of secondary analysis (CRIUGM).

The ADNI2 data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). ADNI was launched in 2003 by the National Institute on Aging, the National Institute

of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership representing efforts of co-investigators from numerous academic institutions and private corporations. ADNI was followed by ADNI-GO and ADNI-2 that included newer techniques. Subjects included in this study were recruited by ADNI-2 from all 13 sites that acquired resting-state fMRI on Philips scanners across North America. For up-to-date information, see www.adni-info.org.

The combined sample included 112 CN and 143 aMCI prior to quality control. After quality control, 99 CN and 129 aMCI remained. In the CN group, the mean age was 72.0 (s.d. 7.0) years, and 37% were men. Mean age of the aMCI subjects was 72.3 (s.d. 7.6) years, and 50% were men. An independent samples t-test did not reveal any significant difference in age between the groups ($t = 0.759, p = 0.448$). A chi-squared test revealed a trend towards a significant difference in gender distribution between the groups ($\chi^2 = 3.627, p = 0.057$). Note that both age and gender were entered as confounding variables in the statistical analysis below. See Table 3.1 for sample size and demographic information from the individual studies after passing quality control (for information about the original cohorts before quality control, see Table 3.4 in supplementary material).

Table 3.1: Demographic information in all studies after quality control

		ADNI2	CRIUGMa	CRIUGMb	MNI	Combined
CN	N	49	18	17	15	99
	Mean age (s.d.)	74.4 (6.8)	71.2 (8.0)	70.4 (4.6)	67.0 (5.7)	72.0 (7.0)
	N male (%)	21 (43%)	7 (39%)	2 (12%)	7 (47%)	37 (37%)
	Mean education (s.d.)[a]	16.9 (2.2)	14.9 (2.3)	15.1 (2.8)	15.0 (3.1)	16.0 (2.6)
	MMSE mean (range)	28.7 (25-30)	28.8 (27-30)	n/a	29.0 (27-30)	n/a
	MoCA mean (range)	n/a	27.8 (22-30)	28.4 (26-30)	n/a	n/a
aMCI	N	82	8	21	18	129
	Mean age (s.d.)	71.2 (7.3)	79.9 (6.1)	74.8 (7.0)	71.2 (8.1)	72.3 (7.6)
	N male (%)	43 (52%)	3 (38%)	12 (57%)	7 (39%)	65 (50%)
	Mean education (s.d.)[a]	16.2 (2.6)	13.7 (3.8)	14.8 (4.2)	13.1 (3.1)	15.5 (3.2)
	MMSE mean (range)	28.1 (24-30)*	26.1 (22-29)*	n/a	26.1 (22-30)*	n/a
	MoCA mean (range)	n/a	23.3 (20-29)*	24.6 (16-29)*	n/a	n/a

MMSE: Mini-mental state examination; MoCA: Montreal Cognitive Assessment.
*Significant difference between aMCI and CN (within study) for independent samples t-test at $p \leq 0.05$.
[a]Missing values for years of education for subjects in ADNI2 (1 CN, 1 aMCI), CRIUGMb (2 aMCI), and MNI (3 CN, 6 aMCI).

All subjects underwent cognitive testing (e.g. memory, language, and executive function) (see Table 3.2 for a list of specific tests used in each study). Exclusion criteria common to all studies included: Contraindications to MRI, presence or history of axis I psychiatric disorders (e.g. depression, bipolar disorder, schizophrenia), presence or history of neurologic disease with potential impact on cognition (e.g. Parkinson's disease), and presence or history of substance abuse. CN subjects could not meet criteria for MCI or dementia. Those with aMCI had memory complaints, objective cognitive loss (based on neuropsychological testing), but had intact functional abilities and did not meet criteria for dementia. In ADNI2, the diagnosis of aMCI was made

based on an education adjusted abnormal score on the Logical Memory II subscale (Delayed Paragraph Recall, Paragraph A only) from the Wechsler Memory Scale and a Clinical Dementia Rating (CDR) of 0.5. In both CRIUGMa and CRIUGMb, the diagnosis of aMCI was made based on scores equal to or greater than 1.5 standard deviations below the mean adjusted for age and education on memory tests. At the MNI, the diagnosis of aMCI relied on the Petersen criteria (Petersen, 2004). At both CRIUGMb and MNI, aMCI diagnoses were made with input from a neurologist. See the Supplementary Methods for greater details for each study.

Table 3.2: Neuropsychological tests that were used in each study

Test	ADNI2	CRIUGMa	CRIUGMb	MNI
Mini-mental State Examination (MMSE)	x	x		x
Montreal Cognitive Assessment (MOCA)	x	x	x	
Clinical Dementia Rating (CDR)	x		x	
ADAS-Cog	x			
Everyday Cognition (ECog)	x			
Trail making	x (Trails A and B)	x (Trails A and B)	x (Trails A and B)	x (DKEFS)
Boston Naming Test	x	x	x	x
Digit span		x	x	x
Colour-word interference (DKEFS)		x	x	x
Rey Auditory Verbal Learning Test	x	x		x
Verbal fluency	x	x	x (MEC)	x (DKEFS)
Clock drawing	x	x		
Visual Object and Space Perception Battery		x		
Brixton Spatial Anticipation Test			x	
Hooper Visual Organization Test			x	
Rey Complex Figure		x	x	x
Aggie Figures Learning Test				x
16-item Free and Cued Recall (RL/RI-16)			x	
Pyramid and Palm Trees Test		x		
Weschler Memory Scale - logical memory subtest	x	x	x	

MEC: Montréal évaluation de la communication; DKEFS: Delis-Kaplan Executive Function System.

3.4.2 Imaging data acquisition

All resting-state fMRI and structural scans were acquired on 3T scanners. We performed analyses on the first usable scan (typically the baseline scan) from ADNI2 and applied clinical diagnoses from the same study time point as the first usable scan for each participant in that dataset. See Table 3.3 for acquisition parameters for each sample.

3.4.3 Computational environment

All experiments were performed using the NeuroImaging Analysis Kit (NIAK[2]) (Bellec et al., 2011) version 0.12.18, under CentOS version 6.3 with Octave[3] version

[2] http://simexp.github.io/niak/
[3] http://gnu.octave.org

Table 3.3: Structural and functional scan acquisition parameters

		ADNI2[a]	CRIUGMa	CRIUGMb	MNI
	Scanner manufacturer	Philips	Siemens	Siemens	Siemens
Structural	No. channels	8	32	32	32
	No. slices	170	176	176	176
	Slice thickness (mm)	1.2	1	1	1
	In-plane resolution (mm x mm)	1 x 1	1 x 1	1 x 1	1 x 1
	Matrix size	256 x 256	240 x 256	256 x 256	256 x 256
	FOV (mm^2)	256	240/256	256	256
	TR (s)	6.8	2.3	2.53	2.3
	TE(ms)	3.09	2.91	1.64	2.98
	TI (s)	n/a	0.9	1.2	0.9
	FA (°)	9	9	7	9
	Slice gap	0	0	0	0
	Imaging plane	Sagittal	Sagittal	Sagittal	Sagittal
	NEX	1	1	1	1
Functional	No. runs	1	1	3	3
	No. channels	8	32	32	32
	No. volumes	140	240	150	160
	No. slices	48	33	42	38
	Slice thickness (mm)	3.3.	4	3.4	3.6
	In-plane resolution (mm x mm)	3.3 x 3.3	3 x 3	3.4 x 3.4	3.6 x 3.6
	Matrix size	64 x 64	64 x 64	64 x 64	64 x 64
	FOV (mm^2)	212	192	218	230
	TR (s)	3	2	2.6	2
	TE (ms)	30	30	30	30
	FA (°)	80	90	90	90
	Slice gap	0	0	0	0
	Imaging plane	Axial	Axial	Axial	Axial
	NEX	1	1	1	1
	Total scan time (min:s)	7:00	8:00	19:30	16:00

[a] http://adni.loni.usc.edu/wp-content/uploads/2011/04/ADNI_3T_Philips_2.6.pdf.

3.8.1 and the Minc toolkit[4] version 0.3.18. Analyses were executed in parallel on the "Guillimin" supercomputer[5], using the pipeline system for Octave and Matlab (Bellec et al., 2012), version 1.0.2. The scripts used for processing can be found on Github[6].

3.4.4 Pre-processing

Each fMRI dataset was corrected for slice timing; a rigid-body motion was then estimated for each time frame, both within and between runs, as well as between one fMRI run and the T1 scan for each subject (Collins and Evans, 1997). The T1 scan was itself non-linearly co-registered to the Montreal Neurological Institute (MNI) ICBM152 stereotaxic symmetric template (Fonov et al., 2011), using the CIVET pipeline (Ad-Dab'bagh et al., 2006). The rigid-body, fMRI-to-T1 and T1-to-stereotaxic transformations were all combined to resample the fMRI in MNI space at a 3 mm isotropic resolution. To minimize artifacts due to excessive motion, all time frames showing a displacement greater than 0.5 mm were removed (Power et al., 2012). A minimum of 50 unscrubbed volumes per run was required for further analysis (13 CN and 14 aMCI were rejected from the original cohort of 112 CN and 143 aMCI).

[4] http://www.bic.mni.mcgill.ca/ServicesSoftware/ServicesSoftwareMincToolKit
[5] http://www.calculquebec.ca/en/resources/compute-servers/guillimin
[6] https://github.com/SIMEXP/mcinet

Neither the rate of rejection nor the frame displacement values (before and after scrubbing) varied significantly among the four samples or between CN and aMCI. The following nuisance covariates were regressed out from fMRI time series: slow time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first 3 to 10 principal components (median numbers for ADNI2, CRIUGMa, CRIUGMb, and MNI were 9, 6, 7, and 7 respectively, and accounting for 95% variance) of the six rigid-body motion parameters and their squares (Giove et al., 2009; Lund et al., 2006). The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel. A more detailed description of the pipeline can be found on the NIAK website[7] and Github[8].

3.4.5 Bootstrap analysis of stable clusters (BASC)

We applied a BASC to identify clusters that consistently exhibited similar spontaneous BOLD fluctuations in individual subjects, and were spatially stable across subjects. We first applied a region-growing algorithm to reduce each fMRI dataset into a time x space array, with 957 regions (Bellec et al., 2006). BASC replicates a hierarchical Ward clustering 1000 times and computes the probability that a pair of regions fall in the same cluster, a measure called stability. The region x region stability matrix is fed into a clustering procedure to derive consensus clusters, which are composed of regions with a high average probability of being assigned to the same cluster across all replications. At the individual level, the clustering was applied to the similarity of regional time series, which was replicated using a circular block bootstrap. Consensus clustering was applied to the average individual stability matrix to identify group clusters. The group clustering was replicated via bootstrapping of subjects in the group. A consensus clustering was finally applied on the group stability matrix to generate group consensus clusters.

The cluster procedure was carried out at a specific number of clusters (called resolution). Using a "multiscale stepwise selection" (MSTEPS) method (Bellec, 2013), we determined a subset of resolutions that provided an accurate summary of the group stability matrices generated over a fine grid of resolutions: 4, 6, 12, 22, 33, 65, 111 and 208.

[7] http://niak.simexp-lab.org/pipe_preprocessing.html
[8] https://github.com/SIMEXP/mcinet/tree/master/preprocess

3.4.6 Derivation of functional connectomes

For each resolution K, and each pair of distinct clusters, the between-clusters connectivity was measured by the Fisher transform of the Pearson's correlation between the average time series of the clusters. The within-cluster connectivity was the Fisher transform of the average correlation between time series inside the cluster. An individual connectome was thus a $K \times K$ matrix. See Figure 3.1a-b for an illustration of a parcellation and associated connectome.

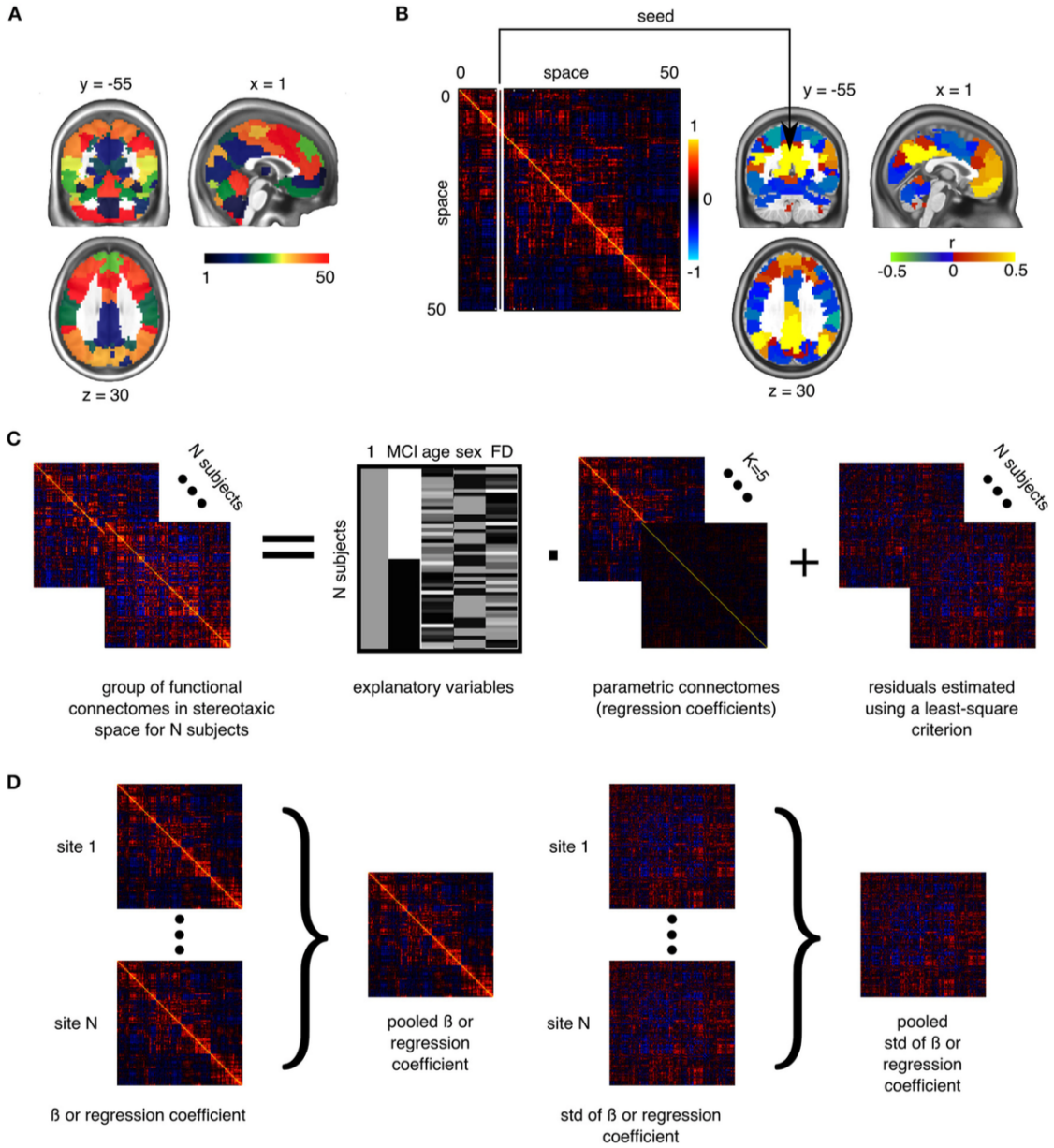

Figure 3.1: Application of general linear models to connectomes. (A) The brain is functionally parcellated into K (e.g. 50) clusters generated through a clustering algorithm. (B) The connectome is a $K \times K$ matrix measuring functional connectivity between and within clusters. (C) A general linear model is used to test the association between phenotypes and connectomes, independently at each connection, at the group level. (D) In a multisite situation, independent site-specific effects are estimated and then pooled through weighted averaging (Willer, Li, and Abecasis, 2010).

3.4.7 Statistical testing

To test for differences between aMCI and CN at a given resolution, we used a general linear model (GLM) for each connection between two clusters. The GLM included an intercept, the age and sex of participants, and the average frame displacement of the runs involved in the analysis. The contrast of interest (aMCI – CN) was represented by a dummy covariate coding the difference in average connectivity between the two groups. All covariates except the intercept were corrected to a zero mean (Figure 3.1c). The GLM was estimated independently for each scanning protocol. In addition to distinguishing between CRIUGMa, CRIUGMb, MNI, and ADNI2, ADNI2 was subdivided into five sub-studies based on the use of different Philips scanner models (i.e. Achieva, Gemini, Ingenia, Ingenuity, and Intera). We dropped all subjects scanned with Ingenuity (2 CN, 1 aMCI) due to the elimination of all aMCI subjects within that site by the scrubbing procedure and its small sample size. We therefore estimated 7 independent GLMs for each protocol (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI). The estimated effects were combined across all protocols through inverse variance based weighted averaging (Willer, Li, and Abecasis, 2010) (Figure 3.1d).

Resolutions containing fewer than 50 clusters have been suggested to have higher sensitivity based on prior independent work (Bellec et al., 2015). The GLM was first applied at an a priori resolution of $K = 33$, which was the lowest number of clusters for which the default mode network could be clearly decomposed into subnetworks (Supplementary Figure 3.7, visit Figshare for 3D volumes of brain parcellations and see Supplementary Table 3.5 for a list of the 33 clusters and their numerical IDs). The false-discovery rate (FDR) across connections was controlled at $q^{\text{FDR}} \leq 0.1$ (Benjamini and Hochberg, 1995). In addition to the analysis at resolution 33, we assessed the impact of that parameter by replicating the GLM analysis at the 7 resolutions selected by MSTEPS (Supplementary Figure 3.8). We implemented an omnibus test (family-wise error rate $\alpha \leq 0.05$) to assess the overall presence of significant differences between groups, pooling FDR results across all resolutions (Bellec et al., 2015). If the omnibus test across resolutions was not significant, then no test would be deemed significant. Since this omnibus test was significant, we used the FDR threshold of $q \leq 0.1$ to explore single resolutions.

3.5 Results

3.5.1 Functional connectivity differences between aMCI and CN

The omnibus test pooling significant differences in connectivity between aMCI and CN across all resolutions was significant at $\alpha \leq 0.05$ ($p = 0.0056$). In line with prior observations on independent datasets (Bellec et al., 2015), resolutions containing fewer than 50 clusters were associated with a higher rate of discovery (Figure 3.2). At resolution 33, significant group differences between aMCI and CN were seen across the whole brain (Figure 3.3a). Four brain clusters were associated with 47% of all significant changes found across the connectome: the superior medial frontal cortex (including anterior cingulate), dorsomedial prefrontal cortex, striatum, and middle temporal lobe (Figures 3.3b, 3.3c, Supplementary Table 3.6). Supplementary Table 3.6 contains a list of parcels that account for all non-redundant significant connectivity differences between aMCI and CN. For example, the first-ranked seed (superior medial frontal cortex) was associated with 13.4% of connections that differ between the groups. The second-ranked seed (dorsomedial prefrontal cortex) was associated with an additional 12.7% of connectivity differences that did not overlap with or were not previously accounted for by the first seed. Note that if a given parcel was associated with a significant effect with another region that ranked in the table, then that parcel may not be listed in the table (i.e. this table is not a comprehensive list of parcels that show significant effects, as a given parcel may involve a region in the table at a higher rank which already accounted for its effects). Given that the top four clusters explained nearly half of the findings, they were further characterized in seed-based connectivity analyses, which revealed that aMCI showed decreased connectivity between default mode network nodes and between areas of the cortico-striatal-thalamic loop (Figure 3.4). More specifically, in aMCI compared to CN, the superior medial frontal cortex displayed significantly reduced connectivity with the ventromedial prefrontal cortex, striatum, thalamus, temporal lobes, hippocampus, inferior parietal lobes, and precuneus (Figure 3.4a). aMCI showed reduced connectivity between the dorsomedial prefrontal cortex with temporal lobe regions, ventral frontal areas, thalamus, striatum, and the cuneus (Figure 3.4b). The striatum in aMCI also exhibited decreased connectivity with the sensorimotor cortex, thalamus, and frontal and parietal regions (Figure 3.4c). Lastly, in aMCI, the middle temporal lobe displayed significantly decreased connectivity with the posterior cingulate, precuneus, inferior parietal lobes, hippocampus, and frontal areas (Figure 3.4d).

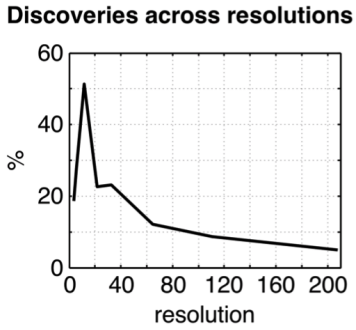

Figure 3.2: Plot of the percentage of connections identified as significant by the statistical comparison between aMCI and CN across the connectome ($q^{\text{FDR}} \leq 0.1$), as a function of the resolutions selected by MSTEPS.

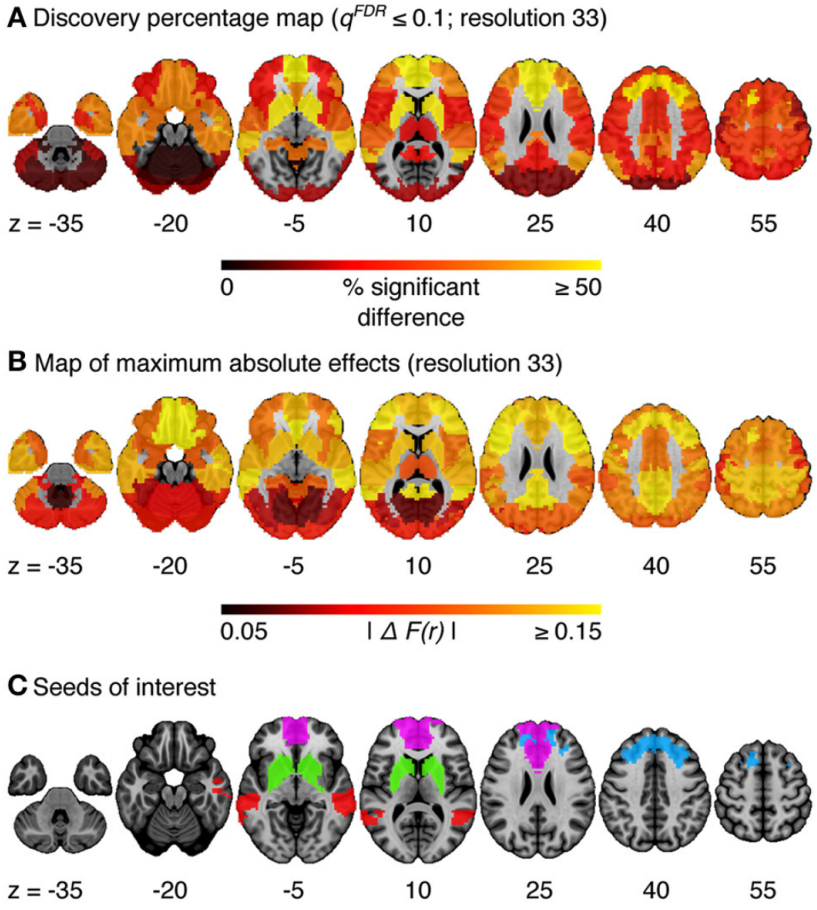

Figure 3.3: (A) Map of the percentage of connections associated with a given cluster and identified as significant by the statistical comparison between aMCI and CN, at a resolution of 33 clusters ($q^{\text{FDR}} \leq 0.1$). (B) Maximum absolute difference in average connectivity between aMCI and CN, across all connections associated with a cluster, at resolution 33. $F(r)$ signifies the difference in Fisher-transformed correlation values between the groups. (C) Four clusters of interest (superior medial frontal cortex, dorsomedial prefrontal cortex, striatum, middle temporal lobe) were selected out of 33 for further characterization.

Figure 3.4: Effect maps for a selection of four seeds that show effects related to aMCI at resolution 33. Effect maps reveal the spatial distribution of the changes in functional connectivity for (A) the superior medial frontal cortex, (B) the dorsomedial prefrontal cortex, (C) striatum, and (D) the middle temporal lobe. All connections shown in the maps of difference in average connectivity between aMCI and CN are significant at $q^{\mathrm{FDR}} \leq 0.1$. For each panel, the top line maps the spatial location of the seed region in magenta, the second and third lines show the connectivity (Fisher-transformed correlation values, $F(r)$) between the designated seed region and the rest of the brain in CN and aMCI, respectively, and the fourth line shows a difference map between aMCI and CN (difference in Fisher-transformed correlation values, $F(r)$). The numbers in parentheses refer to the numerical IDs of the clusters in the 3D parcellation volume, as listed in Supplementary Table 3.5.

3.5.2 Sample-specific effects

The statistical model we used to combine GLM analyses across sites was based on a weighted average. The possibility thus existed that an effect would be significant in the pooled analysis because it was driven by a very strong effect in a single sample, instead of being consistent across all samples. When we examined effects in each sample independently, we detected no findings or very few significant findings. We then explored the whole brain connectivity of the top four seed regions (superior medial frontal cortex, dorsomedial prefrontal cortex, striatum, and middle temporal lobe) within each sample. The majority of effects found at each sample did not appear to be consistent or reproducible across studies as the comparison between aMCI and CN varied substantially among the seven samples (Figure 3.5, Supplementary Figures 3.9 - 3.11). We assessed the extent at which findings among the seven samples were similar by calculating correlation coefficients across the spatial maps for the average connectivity values in CN, the average connectivity values in aMCI, and differences in connectivity values between aMCI and CN among the samples. We found that the difference maps, contrasting aMCI and CN, were weakly correlated on average across studies and protocols (mean $r = 0.06$, min $r = -0.64$, max $r = 0.69$). The average connectivity maps among studies in both CN and aMCI were generally highly correlated with each other (for CN, mean $r = 0.68$, min $r = -0.16$, max $r = 0.95$; for aMCI, mean $r = 0.67$, min $r = -0.10$, max $r = 0.97$). These results were expected given the small sample sizes of most independent samples (Kelly et al., 2012), but still sobering as the majority of the literature on aMCI and fMRI has used small sample sizes.

However, despite the large observed variations in the spatial distribution of aMCI vs CN contrasts, there were still clear consistent trends across studies and protocols. We indeed found that aMCI-related connectivity changes that surpassed the FDR threshold in the pooled analysis showed similar trends in the vast majority of samples across seeds and connections, where the independent aMCI samples consistently exhibited decreased connectivity compared to the CN samples (Figures 3.5 and 3.6, Supplementary Figures 3.9 - 3.11). For example, the pooled analysis revealed that, compared to CN, aMCI exhibited significantly reduced connectivity between the superior medial frontal cortex cortex (the region in which connectivity was most affected by aMCI) and the middle temporal lobes. This change appeared to be common to the majority of the independent samples (Figure 3.5 and Figure 3.6a). For this particular seed, the change in connectivity was mainly due to regions with positive correlations in CN having smaller correlation values closer to zero in aMCI in the individual samples (Figure 3.5 and Figure 3.6a). For sample-specific effects in

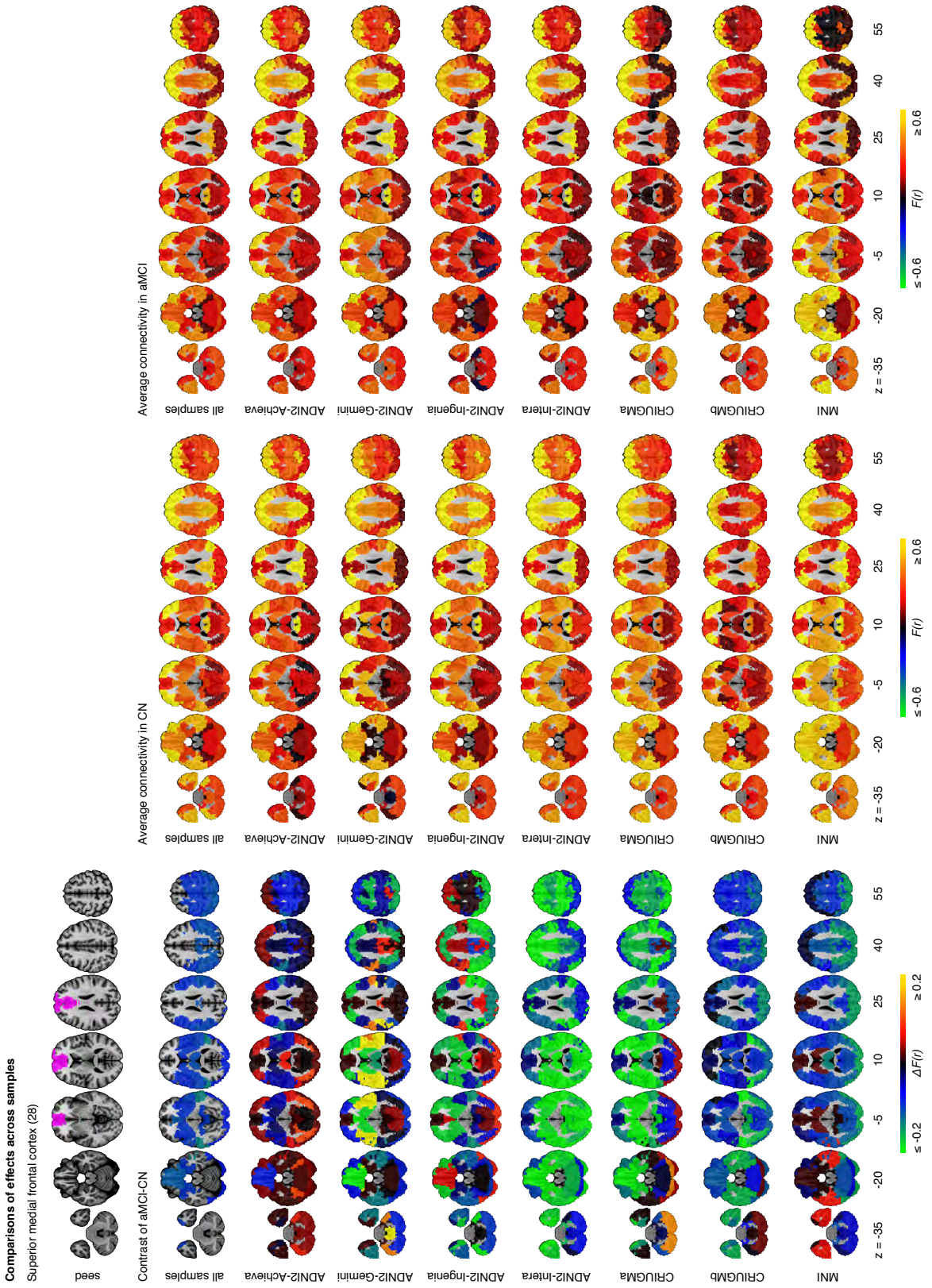

Figure 3.5: Comparisons of effects in the superior medial frontal cortex across samples. This figure illustrates functional connectivity changes between aMCI and CN, average connectivity in CN, and average connectivity in aMCI in each site (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI) independently of other sites and when samples are pooled together (all samples). The number in parentheses refers to the numerical ID of the seed in the 3D parcellation volume, as listed in Supplementary Table 3.5.

other seeds and connections, please see Supplementary Figures 3.9 - 3.15.

3.5.3 Effect sizes and sample size estimates

We measured the effect sizes of the difference between groups at each significant connection by calculating Cohen's d, via a weighted average of the effect sizes per individual sample. We found small to medium effect sizes, ranging from $d = 0.10$ to $d = 0.48$, with an average effect size of $d = 0.32$. Note that these effect sizes are potentially inflated since we have focussed on significant results only. We also calculated the sample sizes required to achieve 80% power, based on the effect sizes estimated by Cohen's d, the assumption of balanced groups, Gaussian distributions, bilateral tests, and $\alpha = 0.05$, for each connection. We found that the estimated sample sizes ranged from 140 to upwards of 600 total subjects, which further suggests that findings from small samples, similar to the seven samples we included when assessed independently, are not expected to be reliable. As noted above, as we used the same sample to estimate the location of effects and their size, these sample size estimates are possibly optimistic, i.e. deflated compared to a replication on an independent sample. See Figure 3.6 and Supplementary Figures 3.12 - 3.15 for Cohen's d and sample size estimates for each significant connection that was reported in Figure 3.4.

3.5.4 Effect of resolution on the GLM

The percentage of discoveries in significant differences between aMCI and CN across the connectomes varied markedly as a function of resolution, as selected by the MSTEPS procedure. Higher resolutions were associated with fewer discoveries, especially beyond resolution 65 (Supplementary Figure 3.16a). By contrast, the maximal amplitude of differences in average connectivity associated with a particular cluster did not decrease substantially, and sometimes increased, when the resolution increased (Supplementary Figure 3.16b). The decrease in percentage of discovery thus likely reflected a cost associated with an increased number of multiple comparisons in the FDR procedure, rather than a loss in signal quality. Regarding the clusters that were selected for our seed-based analyses (the superior medial frontal cortex, dorsomedial prefrontal cortex, striatum, and middle temporal lobe), the associated effect maps (without statistical threshold) were highly consistent across different resolutions (Supplementary Figures 3.17 and 3.18), with the potential exception of very low resolutions where, for example, a relatively small cluster like the anterior cingulate got merged with a large distributed cortical network. This also replicated a prior study on the effect of multiresolution parcellations on GLM analysis (Bellec et al., 2015).

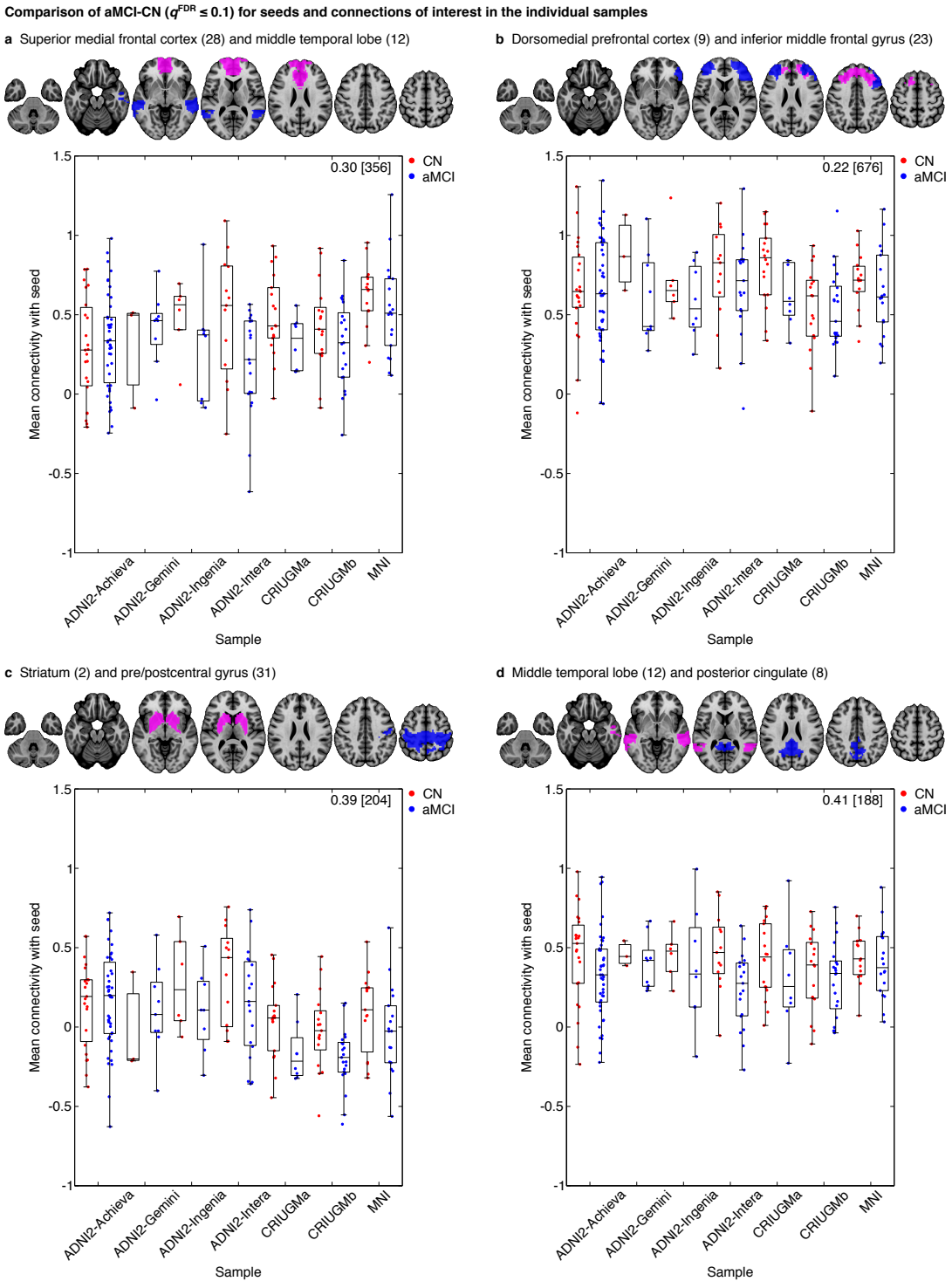

Figure 3.6: Mean connectivity between (A) the superior medial frontal cortex and middle temporal lobe, (B) the dorsomedial prefrontal cortex and middle frontal gyrus, (C) the striatum and pre/postcentral gyrus, and (D) middle temporal lobe and posterior cingulate in CN and aMCI in the independent samples. Each map displays the seed (pink) and a selected cluster (blue) whose connectivity with the seed significantly differed between CN and aMCI in the pooled analysis. The box-whisker plots display the mean connectivity (Fisher-transformed correlation values) between the seed and the selected parcel, overlaid over individual data points, in the CN and MCI groups in the ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, and MNI samples. We also report the Cohen's d (a weighted average of the effect sizes per sample) followed by a sample size estimate (for 80% power, balanced groups, bilateral tests, Gaussian distributions, and $\alpha = 0.05$) in square brackets in the top-right corner of each plot. The numbers in parentheses in the titles refer to the numerical IDs of the seeds in the 3D parcellation volume, as listed in Supplementary Table 3.5. For box-whisker plots for all significant clusters with each of these seeds, see Supplementary Figures 3.12 – 3.15.

3.6 Discussion

We report resting-state functional connectivity differences in the superior medial frontal cortex, dorsomedial prefrontal cortex, striatum, and middle temporal lobe between aMCI and CN subjects when multiple studies were combined together. Despite protocol differences, we found that aMCI exhibited reduced connectivity within areas of the default mode network and cortico-striatal-thalamic loop compared to CN. Previous studies suggested these altered patterns of functional connectivity in MCI may result from the coevolution of multiple AD-associated biological processes, namely structural degeneration (Coupé et al., 2012; Pievani et al., 2010), neurofibrillary and amyloid pathologies (Small et al., 2006), and cerebrovascular dysfunction (Villeneuve and Jagust, 2015).

The superior medial frontal cortex and middle temporal lobes, both of which are default mode network nodes, were among the seed regions with the greatest amount of aMCI-related connectivity changes with other brain areas. Decreased connectivity in aMCI patients was found between these two nodes and other default mode network regions, including the posterior cingulate, precuneus, inferior parietal lobes, ventromedial prefrontal cortex, and hippocampus. Our findings support previous studies that used small single-site samples and reported reduced default mode network connectivity in MCI and AD patients (Agosta et al., 2012; Bai et al., 2009; Koch et al., 2012; Sorg et al., 2007). Alterations in the default mode network may reflect increased amyloid burden in aMCI patients as it has been shown that amyloid plaques impair default mode connectivity (Hedden et al., 2009; Mormino et al., 2011; Sheline et al., 2010a).

We found reduced connectivity within the frontal lobes, notably between ventral and dorsal areas. Decreased functional connectivity between the ventral and dorsal frontal regions could reflect degeneration in gray matter and in white matter tracts connecting these areas. Longitudinal studies have shown greater prefrontal cortex atrophy in MCI over time, as well as in those transitioning to AD, compared to CN (Carmichael et al., 2013; McDonald et al., 2009). Cortico-cortical white matter bundles, e.g. superior longitudinal fasciculus, have also been demonstrated to degenerate in patients with MCI and AD (Pievani et al., 2010). Additionally, functional connectivity changes may reflect the regional effect of increased amyloid burden (Sheline et al., 2010a), and PIB-PET work has shown the frontal lobe to be one of the first regions in which amyloid accumulates in autosomal dominant AD mutation carriers (Bateman et al., 2012). Our results may also be due to neurofibrillary pathology as it typically appears in the prefrontal cortex during MCI (Bossers et al., 2010). Lastly, cerebral hypoperfusion in the frontal lobe of MCI (Chao et al., 2009) may have contributed to

our results.

We also observed functional disconnection between the temporal and frontal lobes in aMCI. Effects in the temporal lobes were expected given that the temporal lobe is a region known to suffer from significant AD pathology in preclinical phases (Guillozet et al., 2003). Structural connectivity may also explain the functional connectivity changes between the frontal and temporal regions, since degeneration of white matter tracts between these areas, e.g. the uncinate fasciculus, occurs with the progression from MCI to AD and correlates with episodic memory impairment in MCI (Pievani et al., 2010; Rémy et al., 2015). Furthermore, examining the integrity of the arcuate fasciculus, a major language tract that connects the frontal and temporal lobes (Dick and Tremblay, 2012), might reveal a biological basis for language impairments such as word-finding difficulties in MCI and AD (Nutter-Upham et al., 2008). Brain areas that subserve language function could be important targets to investigate given recent evidence that multilingualism, like other forms of cognitive reserve, may help delay the onset of AD (Chertkow et al., 2010).

Unexpectedly, we also found significant effects in the striatum, which showed reduced connectivity in aMCI with sensorimotor cortex, frontal and parietal regions, and thalamus. While not initially expected, these findings may reflect earlier observations that regions within the cortico-striatal-thalamic loops are vulnerable to AD pathology. For example, previous work demonstrated the presence of substantial amyloid burden in the striatum in both autosomal dominant and sporadic forms of AD (Braak and Braak, 1990; Villemagne et al., 2009), and the striatum may be the first region in which amyloid deposition occurs in autosomal dominant AD (Bateman et al., 2012; Klunk et al., 2007). Furthermore, significant neurodegeneration is known to occur with AD in the striatum and thalamus (Jong et al., 2008; Madsen et al., 2010), so our results might reflect the brain's capacity for functional plasticity in response to amyloid or neurodegeneration in these regions. Motor cortex hyperexcitability has also been shown in AD, and this suggests that inhibitory circuits leading to the motor cortex may be affected in the disease (Ferreri et al., 2011). Patients with AD also demonstrate changes in swallowing which have been associated with altered cortical activity (Humbert et al., 2010). Our results may support these observations. Additionally, our findings may represent a biological basis for the cognitive and motor symptoms of MCI (Aggarwal et al., 2006) since the striatum and the rest of the basal ganglia have been implicated in stimulus-response associative learning and memory and motor skill acquisition and execution (Doyon et al., 2009; Packard and Knowlton, 2002). Future research should examine the potential relationship between connectivity in the cortico-striatal-thalamic loops and motor function in aMCI and AD.

Our findings contrasted with previous, smaller single-site studies that have variously reported decreased and increased connectivity. The reports of increased connectivity (Bai et al., 2009; Gour et al., 2011; Qi et al., 2010) may have reflected unique attributes of particular protocols or the choices made with respect to preprocessing steps, for example using global signal regression (Saad et al., 2012). Given that our sample size estimates suggest the use of hundreds of subjects to obtain adequate statistical power, it is not surprising that discrepancies between our results and previous findings generated from smaller, likely underpowered, studies exist. Even when we examined the samples in our study (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI) independently of each other, we found inconsistent effects among the samples. It is only by combining the studies together in a meta-analysis that we were able to find some common differences in functional connectomes between patients with aMCI and CN. This finding underscores the need for multisite studies with large sample sizes in order to generate reproducible results, as previously suggested in the field of autism research (Haar et al., 2014).

Among our study's limitations is that it was not possible to model each of the 13 ADNI2 sites independently because the sites tended to be small and unbalanced in the numbers of patients and controls. We therefore chose to model each scanner model within ADNI2 separately based on the recommendation of a reviewer. A previous version of the analysis (published as a preprint[9]) had not modeled the different scanner models in ADNI2 and instead treated ADNI2 as a single site. This previous analysis yielded fewer significant findings, but the results were still mostly consistent with what is reported here. Our results suggest that modeling scanner models may have a positive impact on fMRI association studies, but further experiments would be required to confirm that this trend is reproducible. We must also note that the METAL averaging is only representative of the specific samples that were averaged, especially using only Philips and Siemens scanners, and it is unclear how our findings may replicate in other studies that would employ a different combination of protocols, say using GE scanners. In particular, our sample size estimates have to be interpreted with caution. They may first be under-estimated, because they were not derived from pre-specified locations, but rather associated with the connections showing the largest effects in our particular sample. These sample sizes were also derived from a meta-analysis combining particular types of studies. We only had 3T scanners from two manufacturers, Siemens and Philips. For the Siemens studies, all were from the same model. For the Philips studies, the scanning protocol was identical at every site, and only the scanner model varied across scanners. Finally, a fairly large number of

[9] http://dx.doi.org/10.1101/019646

patients and controls (generally more than 10 subjects per group) was scanned for each variant of the scanning protocol. The sample size estimate may turn out quite differently for a single site study or on the contrary for a study with a very large number of sites and with only a few subjects per site.

Our study is also limited by its cross-sectional nature, which precludes inference that the functional changes we found would necessarily predict progression towards Alzheimer's dementia. Furthermore, aMCI has many underlying causes aside from AD. It is possible that some subjects in our cohort had cognitive impairments due to Lewy Body dementia, for example. However, all samples in the current study had inclusion criteria that enriched for subjects that had aMCI likely due to AD and excluded aMCI subjects with other co-morbidities, such as depression or Parkinson's disease. Also, we did not account for structural atrophy, despite a bias for increased detection in functional differences due to differences in underlying structure (Dukart and Bertolino, 2014). However aMCI-related gray matter changes likely co-localize to some extent with functional changes, and the aim of our work was to map out functional changes rather than study their interaction with atrophy. We did not account for other variables, such as APOE genotype (Sheline et al., 2010b), amyloid deposition (Sheline et al., 2010a), presence of neurofibrillary tangles (Maruyama et al., 2013), and cerebrovascular mechanisms (Villeneuve and Jagust, 2015). At least some of these could potentially have explained the observed aMCI-related functional connectivity changes as part of an underlying disease mechanism. Large-scale multimodal studies, incorporating genomics, proteomics, and multimodal imaging will be needed to identify the interactions between these and other physiological facets of the pathology. Despite combining several samples together, we still only achieved relatively limited power, given that sample size estimates required at least 140 to over 600 total subjects to consistently identify effects between groups. Lastly, because of the explorative approach used in our study, the resulting estimates of effect sizes may have been inflated and discussion of possible pathological mechanisms for our findings was speculative. However, our discoveries may be used as follow-up targets in future work. Upcoming research should not only attempt to verify our findings by using these regions and their associated connections with hypothesis-driven approaches (e.g. seed-based correlation analyses), but also to extend them to cohorts that include Alzheimer's dementia and other clinical populations (e.g. CN with significant amyloid deposition) and to longitudinal studies that characterize individuals' progression to dementia. Finally, future studies should aim to determine whether our findings are associated with established biomarkers of AD (e.g. amyloid and tau quantification) in order to probe the potential of these functional connections as biomarkers.

Overall, our results supported previous findings of DMN connectivity changes in AD and MCI (Greicius et al., 2004; Sorg et al., 2007), given that three of the identified seeds (superior medial frontal cortex, dorsomedial prefrontal cortex, middle temporal lobe) are part of this network. It is noteworthy, however, that our strongest observed effects reported were not in the same DMN regions typically described in earlier resting-state studies of MCI and AD, viz, posterior cingulate/precuneus (Sheline et al., 2010a; Zhang et al., 2010). Unexpected changes were also found in the striatum, and this may reflect the advantages of "mining" the whole-brain connectome to search for new biomarkers of mild cognitive impairment and possibly the early progression of the pathophysiologic substrate of Alzheimer's disease. If confirmed, our results could suggest the utility of these regions in resting-state fMRI as a biomarker endpoint in clinical trials.

3.7 Acknowledgements

We thank Sven Joubert, Isabelle Rouleau, and Sophie Benoit for their contribution in collection of the CRIUGMa dataset. We thank Tom Nichols for his comments and suggestions. Collection of the CRIUGMb dataset was supported by the Alzheimer's Society of Canada. Collection of the MNI dataset was supported by Industry Canada/Montreal Neurological Institute Centre of excellence in commercialization and research. This work was supported by CIHR (133359) and NSERC (436141-2013).

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 DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing 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 coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

3.8 Supplementary material

3.8.1 Supplementary methods

Inclusion/exclusion criteria by each individual study

ADNI2

Subjects must have been either English or Spanish-speaking and between 55-90 (inclusive) years of age. All subjects had a partner able to provide an independent evaluation of functioning. Criteria for CN subjects were as follows: MMSE scores between 24-30 (inclusive), a CDR of 0, non-depressed, non-MCI, and non-demented. Criteria for aMCI were as follows: MMSE scores between 24-30 (inclusive), a memory complaint, have objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia.

CRIUGMa

All subjects must have been 50 years of age or older, French-speaking or bilingual, and lived in Quebec for most of their lives. Exclusion criteria included history of psychiatric or neurological disorders, history of substance abuse, and having undergone anesthesia within the last six months of participating in the study. aMCI must have subjective memory complaints, an objective memory deficit (measured as a score at least 1.5 standard deviations below the mean considering age and education), and intact functional abilities. aMCI must not have met criteria for dementia. CN subjects must not have met criteria for MCI or dementia.

CRIUGMb

All subjects must have been between 65 and 80 years old, had normal audition and vision at the time of study, and at least 8 years of education. Exclusion criteria included family history of early-onset AD, use of psychoactive substances within last 3 months, intellectual disabilities, signs of depression. aMCI must have had memory

complaints, an MMSE score equal to or greater than 24, a CDR of 0.5, and cognitive deficits equal to or higher than 1.5 standard deviations on neuropsychological tests. CN subjects must not have met criteria for MCI or dementia.

MNI

CN subjects were selected on basis of their neurological and clinical status. Initial identification of patients was based on memory complaints substantiated by an informant. A subsequent interview was conducted with a full neurological examination including the standard Mini Mental State Examination (MMSE). Routine blood screening was done to rule out underlying metabolic disorder. aMCI diagnosis was according to the Petersen criteria (Petersen, 2004) which included: (i) memory complaint usually corroborated by an informant; (ii) objective memory impairment for age; (iii) essentially preserved general cognitive function; (iv) largely intact functional activities; (v) not demented. Exclusion criteria included co-morbidity with other neurological disease such as stroke, Parkinson's disease, other neurodegenerative diseases, etc; the presence of any major structural abnormalities or signs of major vascular pathology on the MRI evaluation; axis I psychiatric disorder or intellectual disability; use of psychoactive substance; previous or present use of cholinesterase inhibitor.

3.8.2 Supplementary figures

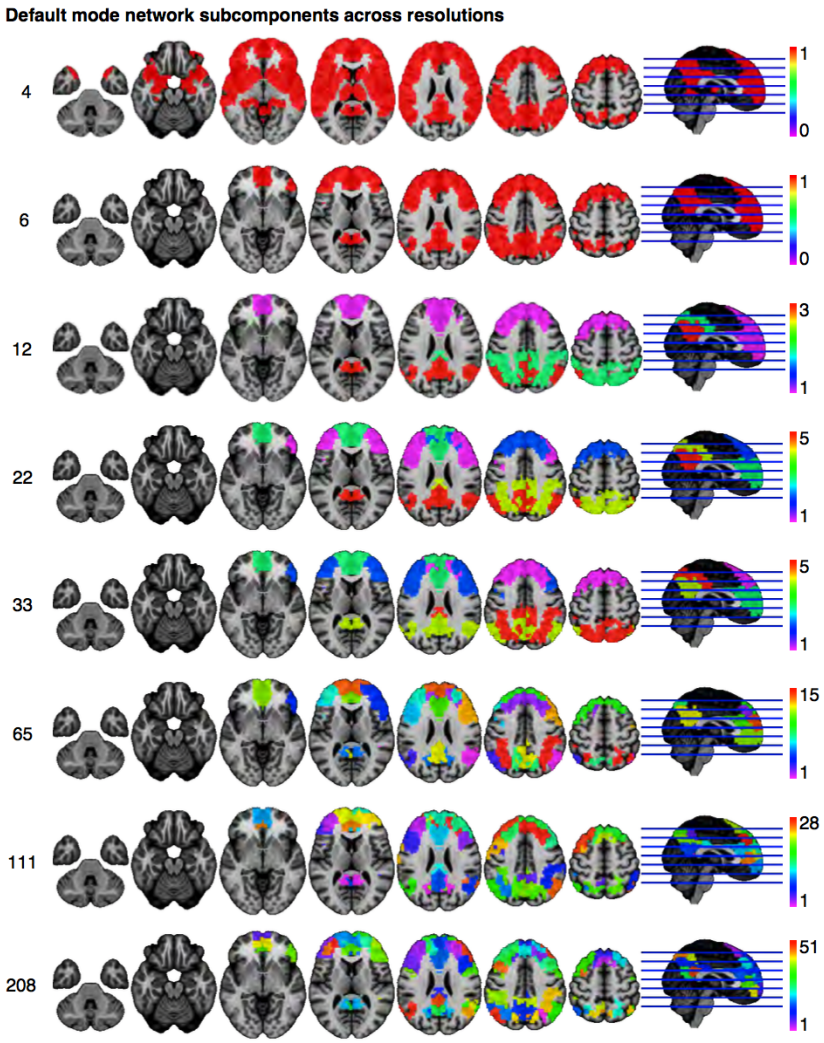

Figure 3.7: The default mode network and its subcomponents across resolutions (or number of clusters) selected by MSTEPS.

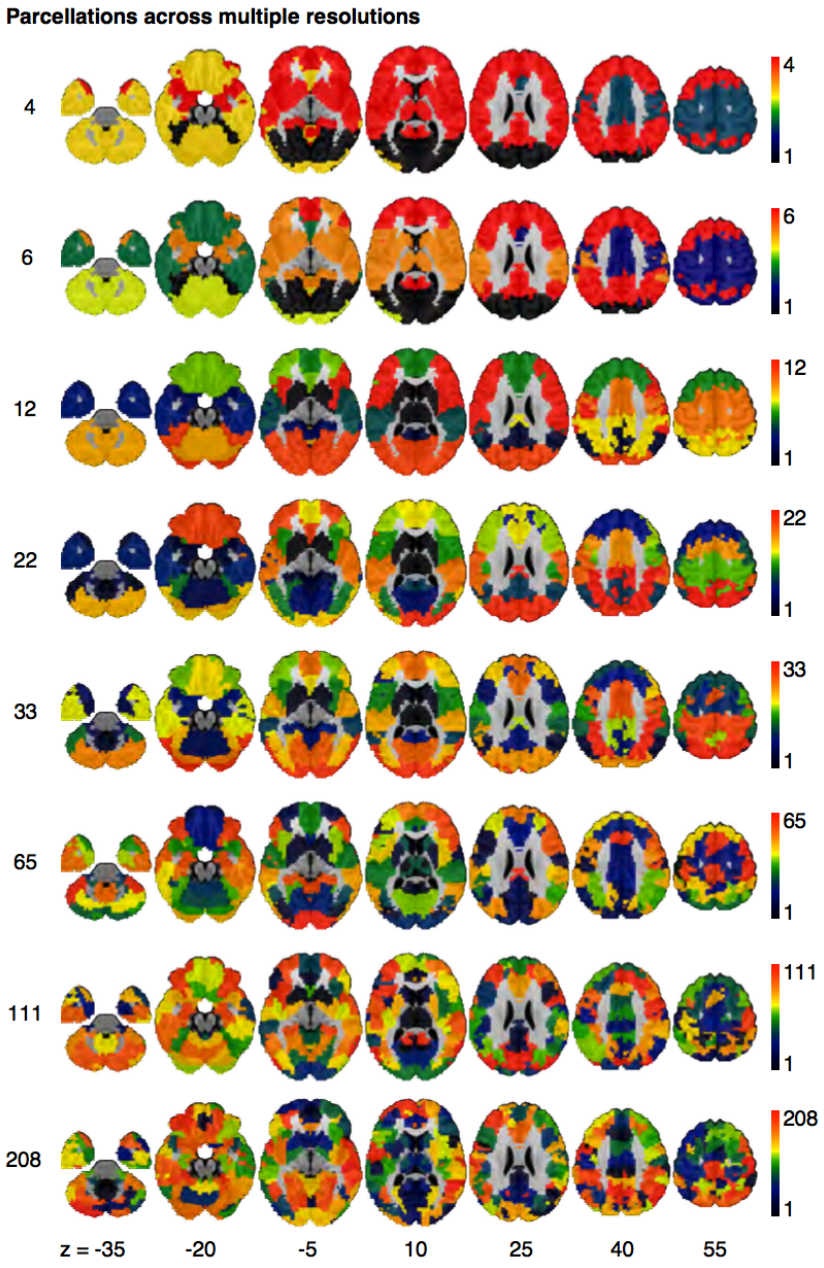

Figure 3.8: Functional parcellations across resolutions (or number of clusters) selected by MSTEPS.

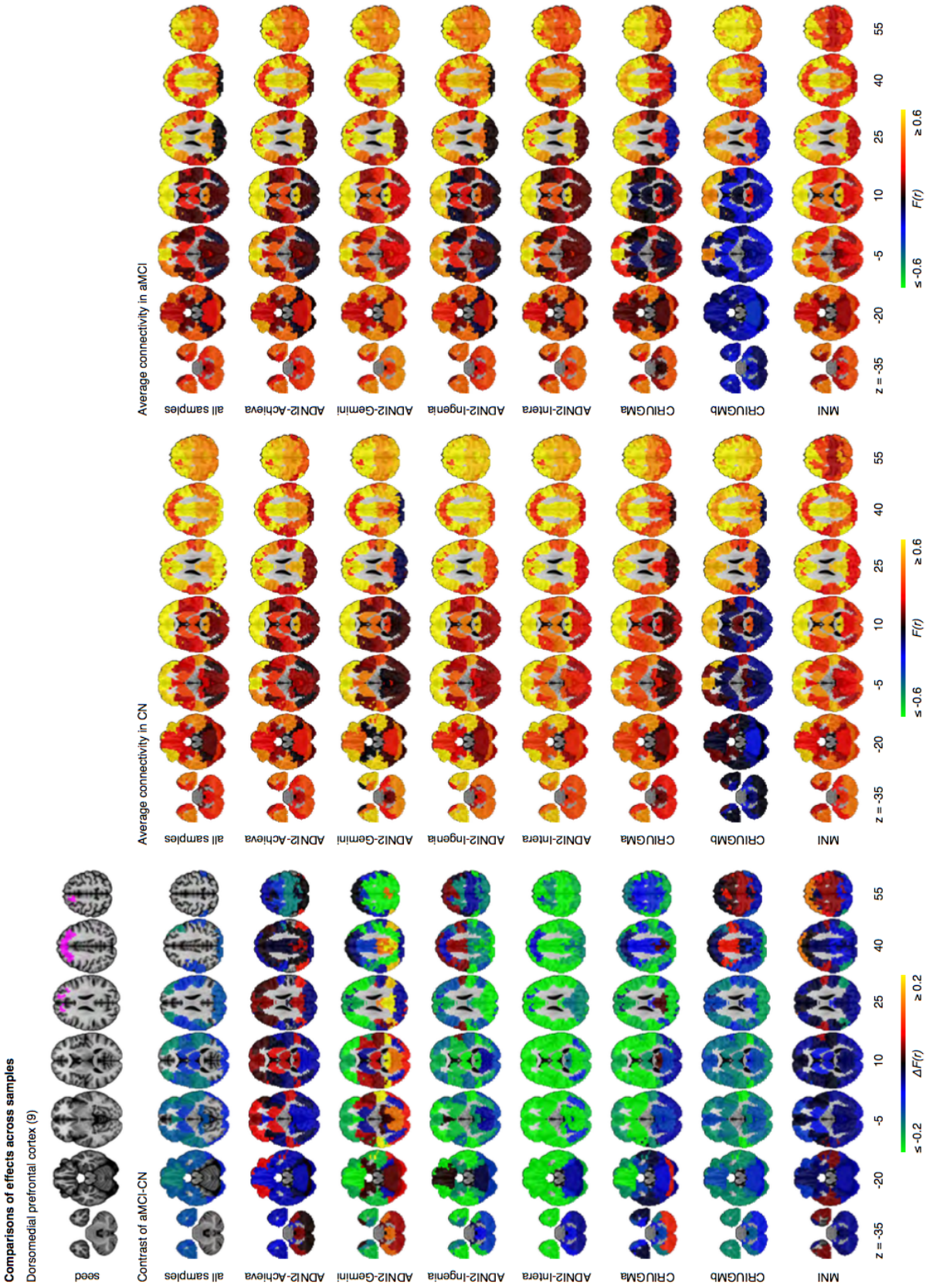

Figure 3.9: Comparisons of effects in the dorsomedial prefrontal cortex across samples. This figure illustrates functional connectivity changes between aMCI and CN, average connectivity in CN, and average connectivity in aMCI in each site (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI) independently of other sites and when samples are pooled together (all samples). The number in parentheses refers to the numerical ID of the seed in the 3D parcellation volume, as listed in Supplementary Table 3.5.

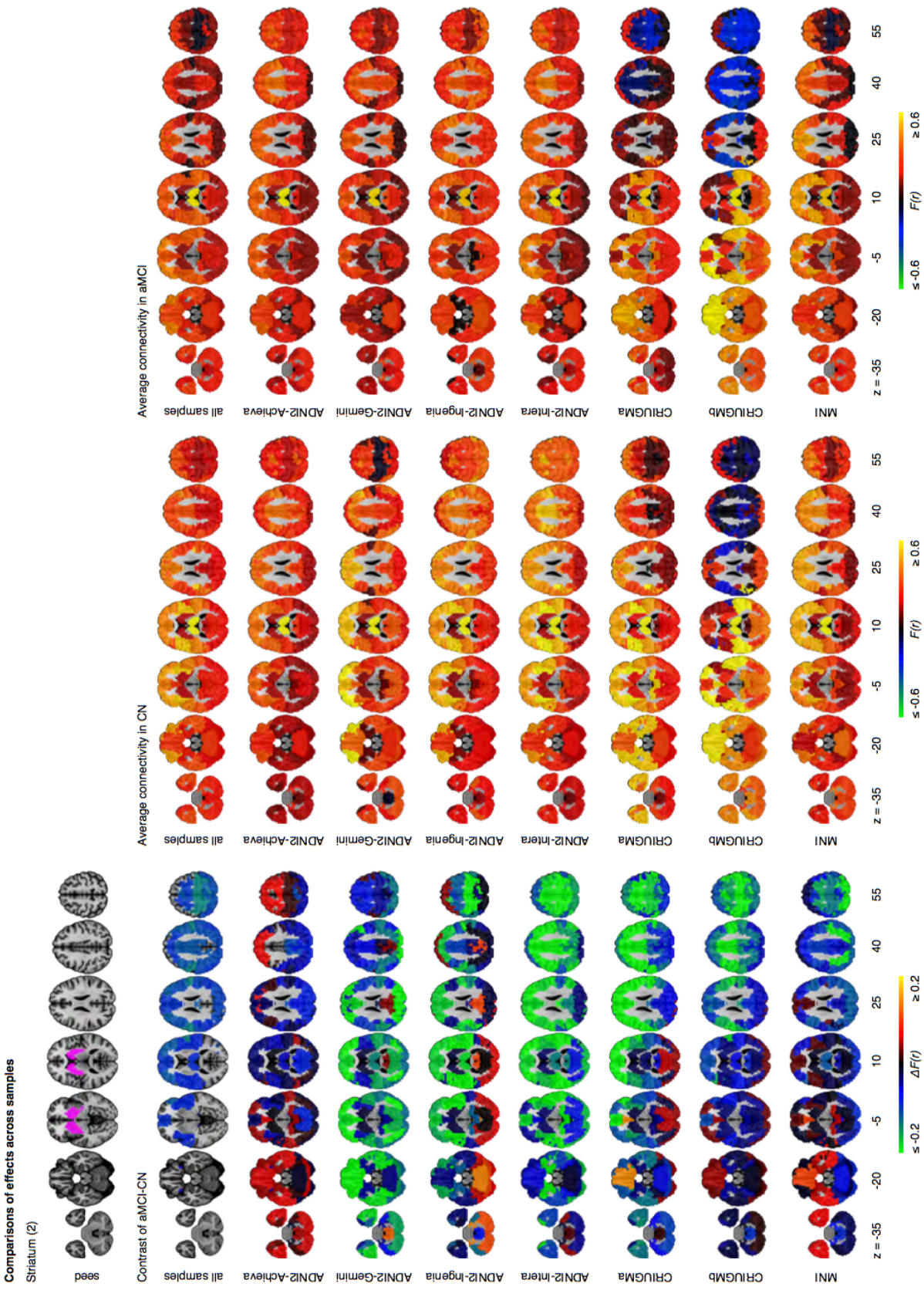

Figure 3.10: Comparisons of effects in the striatum across samples. This figure illustrates functional connectivity changes between aMCI and CN, average connectivity in CN, and average connectivity in aMCI in each site (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI) independently of other sites and when samples are pooled together (all samples). The number in parentheses refers to the numerical ID of the seed in the 3D parcellation volume, as listed in Supplementary Table 3.5.

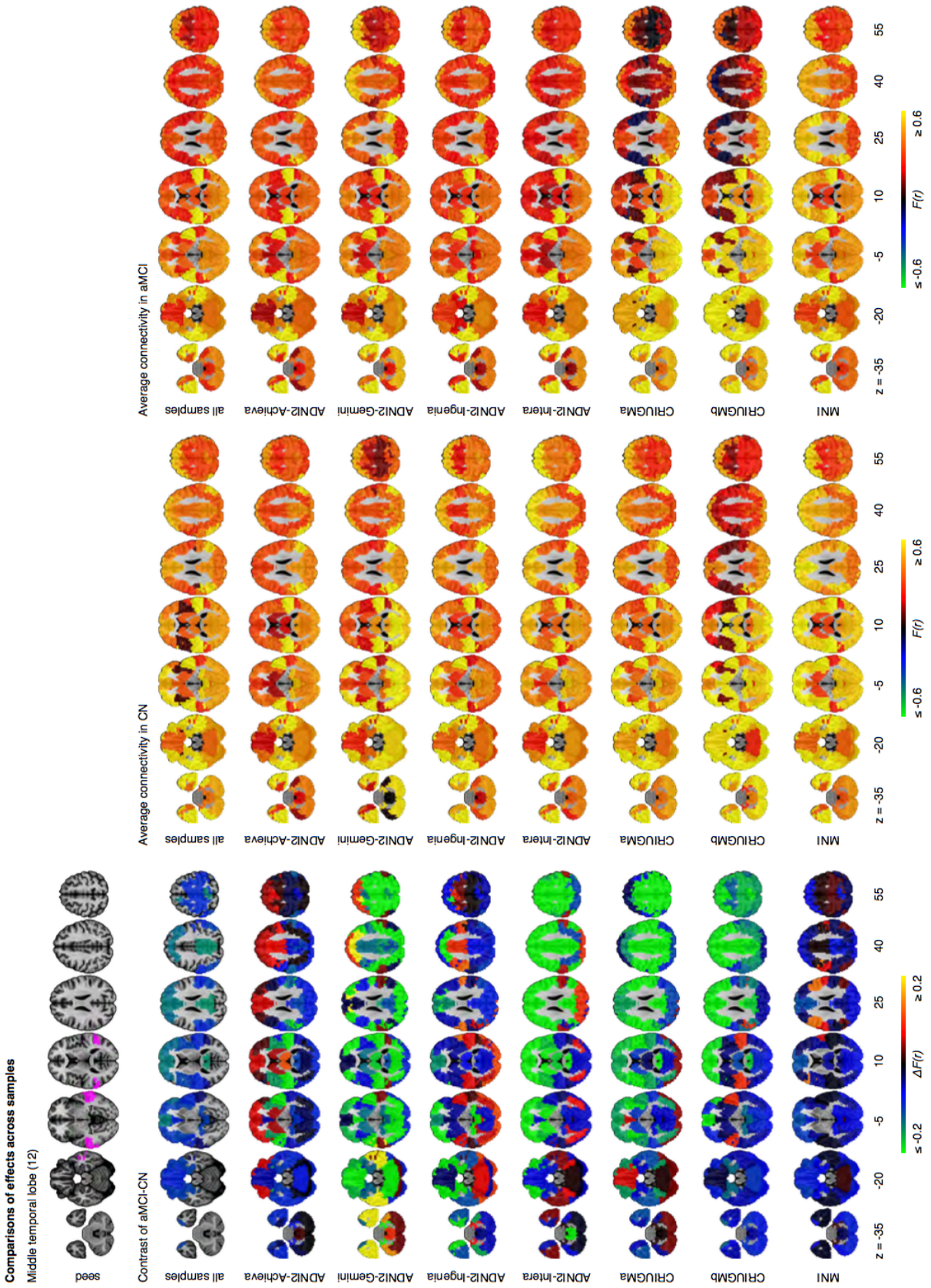

Figure 3.11: Comparisons of effects in the middle temporal lobe across samples. This figure illustrates functional connectivity changes between aMCI and CN, average connectivity in CN, and average connectivity in aMCI in each site (ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, MNI) independently of other sites and when samples are pooled together (all samples). The number in parentheses refers to the numerical ID of the seed in the 3D parcellation volume, as listed in Supplementary Table 3.5.

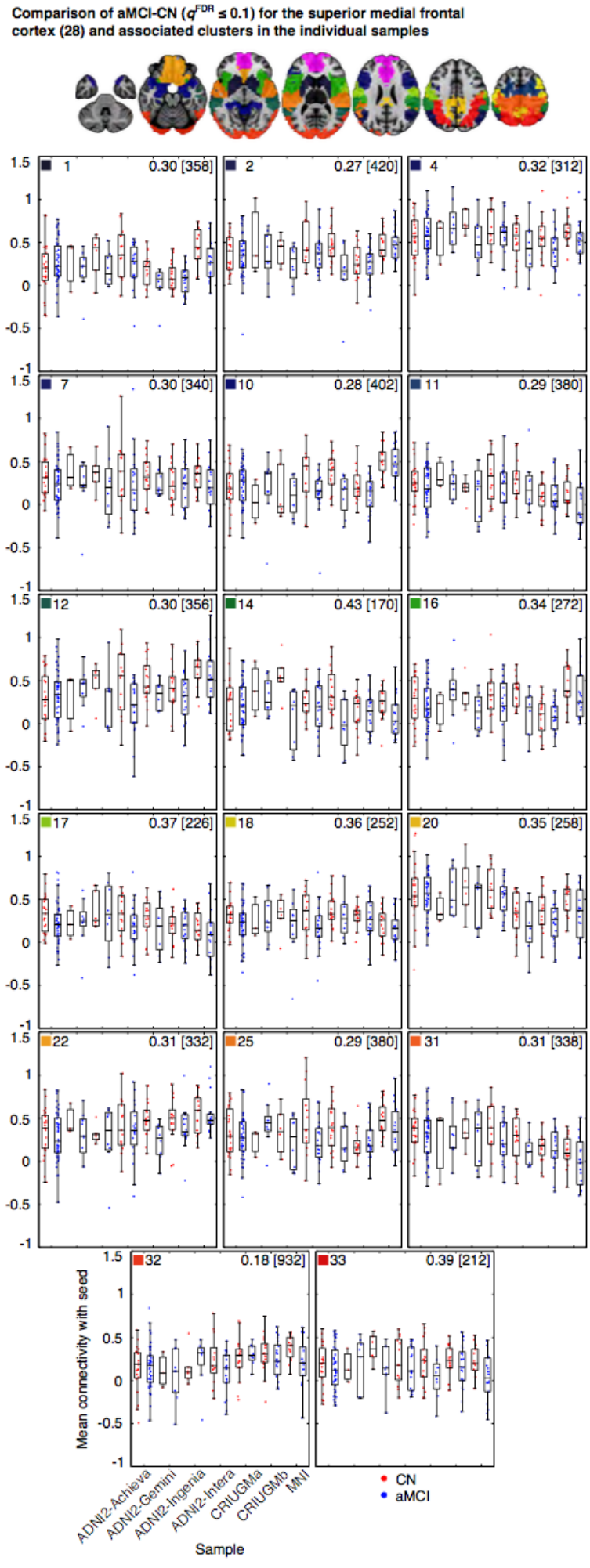

Figure 3.12: Mean connectivity of the superior medial frontal cortex (cluster #28) with its associated connections in CN and aMCI in the independent samples. Each map displays the seed in pink and the clusters (in other colors) whose connectivity with the seed significantly differed between CN and aMCI in the pooled analysis. The box-whisker plots display the mean connectivity (Fisher-transformed correlation values) between the seed and a significant parcel, overlaid over individual data points, in the CN and aMCI groups in the ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, and MNI samples. Each plot is labeled with a number in the top-left corner, corresponding to the number assigned to the cluster in Supplementary Table 3.5, and a colored square corresponding to the parcel of the same color from the map. We also report the Cohen's d (a weighted average of the effect sizes per sample) followed by a sample size estimate (for 80% power, balanced groups, bilateral tests, Gaussian distributions, and $\alpha = 0.05$) in square brackets in the top-right corner of each plot.

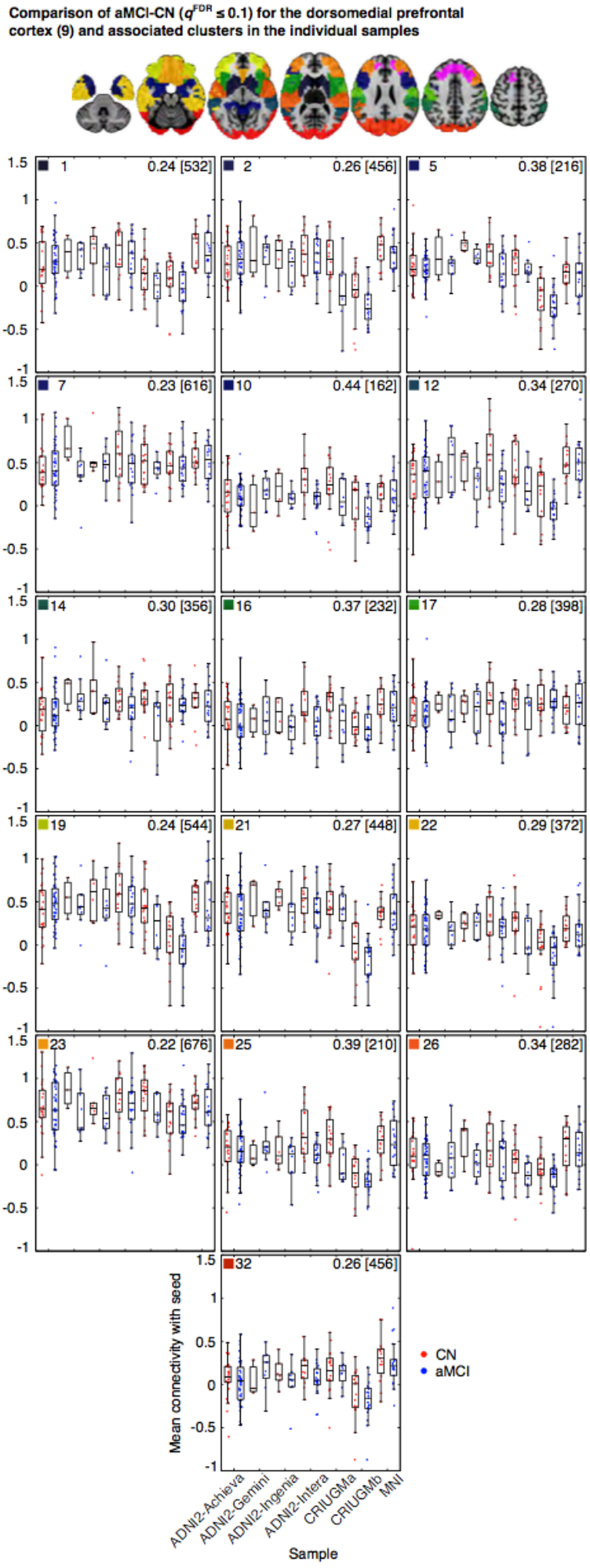

Figure 3.13: Mean connectivity of the dorsomedial prefrontal cortex (cluster #9) with its associated connections in CN and aMCI in the independent samples. Each map displays the seed in pink and the clusters (in other colors) whose connectivity with the seed significantly differed between CN and aMCI in the pooled analysis. The box-whisker plots display the mean connectivity (Fisher-transformed correlation values) between the seed and a significant parcel, overlaid over individual data points, in the CN and aMCI groups in the ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2- Intera, CRIUGMa, CRIUGMb, and MNI samples. Each plot is labeled with a number in the top-left corner, corresponding to the number assigned to the cluster in Supplementary Table 3.5, and a colored square corresponding to the parcel of the same color from the map. We also report the Cohen's d (a weighted average of the effect sizes per sample) followed by a sample size estimate (for 80% power, balanced groups, bilateral tests, Gaussian distributions, and $\alpha = 0.05$) in square brackets in the top-right corner of each plot.

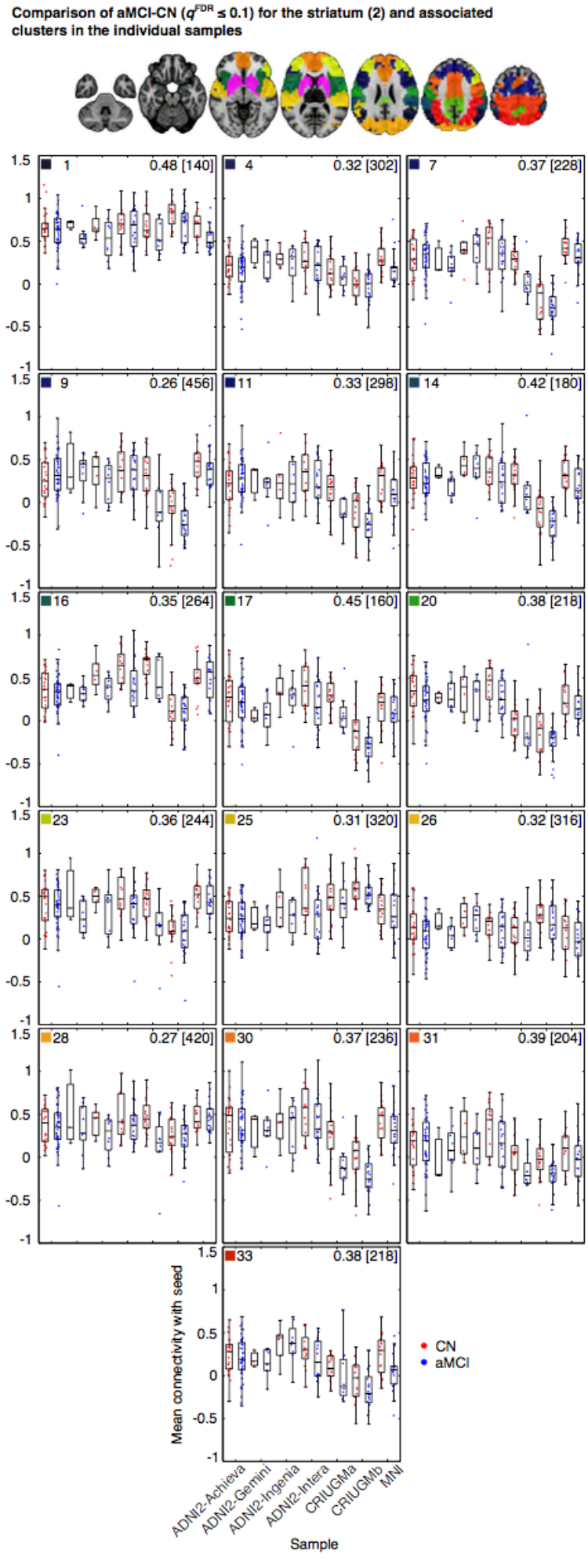

Figure 3.14: Mean connectivity of the striatum (cluster #2) with its associated connections in CN and aMCI in the independent samples. Each map displays the seed in pink and the clusters (in other colors) whose connectivity with the seed significantly differed between CN and aMCI in the pooled analysis. The box-whisker plots display the mean connectivity (Fisher-transformed correlation values) between the seed and a significant parcel, overlaid over individual data points, in the CN and aMCI groups in the ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, and MNI samples. Each plot is labeled with a number in the top-left corner, corresponding to the number assigned to the cluster in Supplementary Table 3.5, and a colored square corresponding to the parcel of the same color from the map. We also report the Cohen's d (a weighted average of the effect sizes per sample) followed by a sample size estimate (for 80% power, balanced groups, bilateral tests, Gaussian distributions, and $\alpha = 0.05$) in square brackets in the top-right corner of each plot.

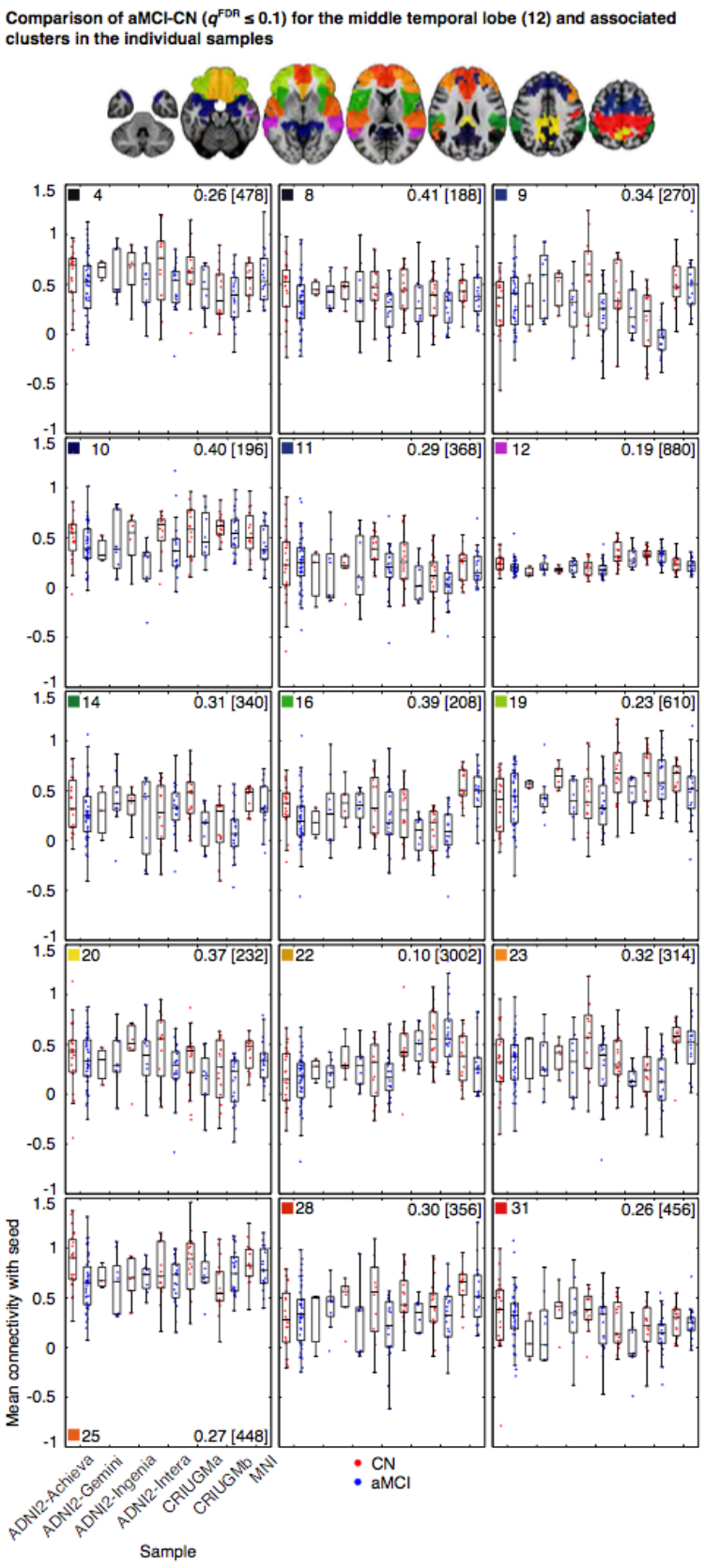

Figure 3.15: Mean connectivity of the middle temporal lobe (cluster #12) with its associated connections in CN and aMCI in the independent samples. Each map displays the seed in pink and the clusters (in other colors) whose connectivity with the seed significantly differed between CN and aMCI in the pooled analysis. The box-whisker plots display the mean connectivity (Fisher-transformed correlation values) between the seed and a significant parcel, overlaid over individual data points, in the CN and aMCI groups in the ADNI2-Achieva, ADNI2-Gemini, ADNI2-Ingenia, ADNI2-Intera, CRIUGMa, CRIUGMb, and MNI samples. Each plot is labeled with a number in the top-left corner, corresponding to the number assigned to the cluster in Supplementary Table 3.5, and a colored square corresponding to the parcel of the same color from the map. We also report the Cohen's d (a weighted average of the effect sizes per sample) followed by a sample size estimate (for 80% power, balanced groups, bilateral tests, Gaussian distributions, and $\alpha = 0.05$) in square brackets in the top-right corner of each plot.

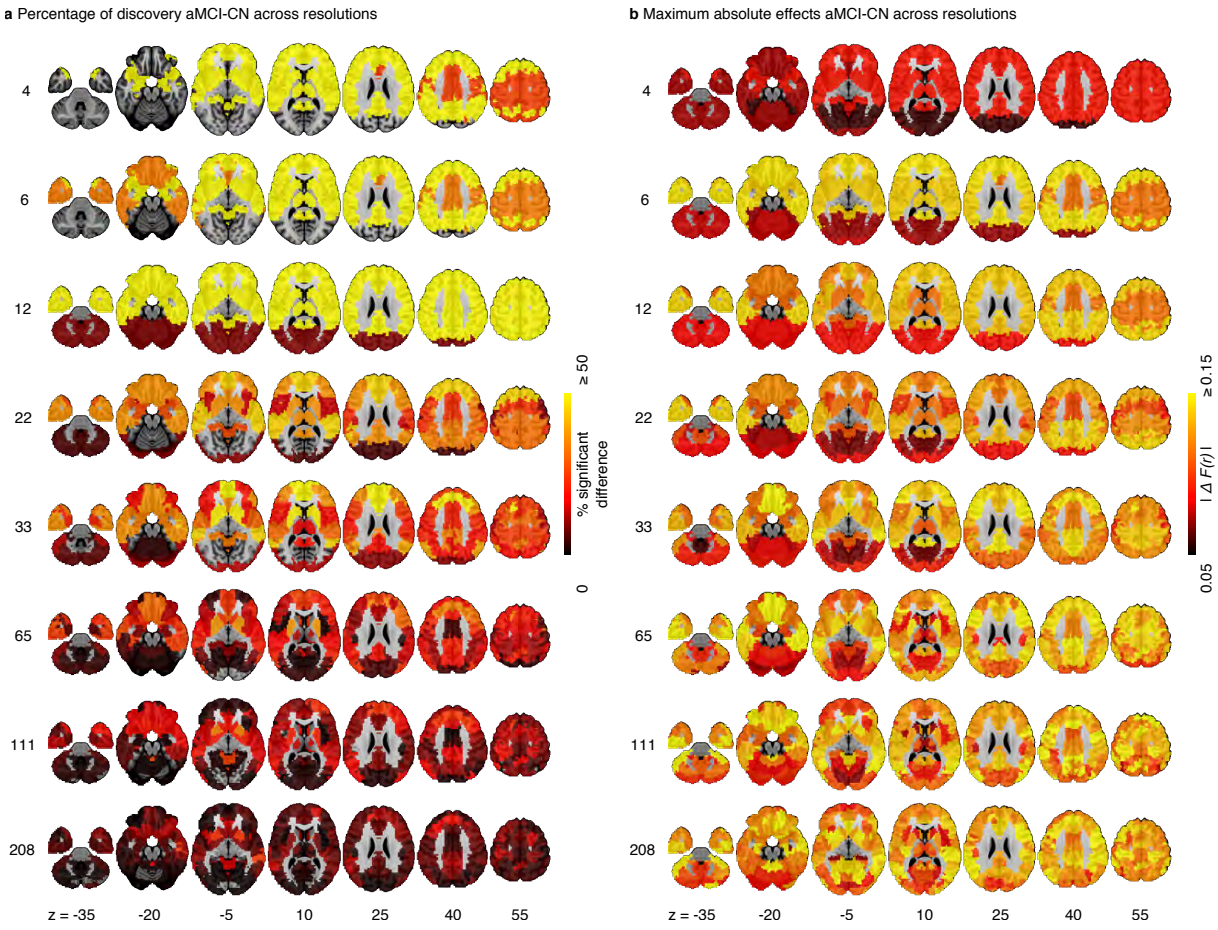

Figure 3.16: Comparison of results across different resolutions (or number of clusters) selected by MSTEPS. (A) Maps of percentage of discovery illustrating brain networks that are significantly different between aMCI and CN. (B) Maps of maximum absolute effects show the magnitude of the differences between aMCI and CN of every cluster.

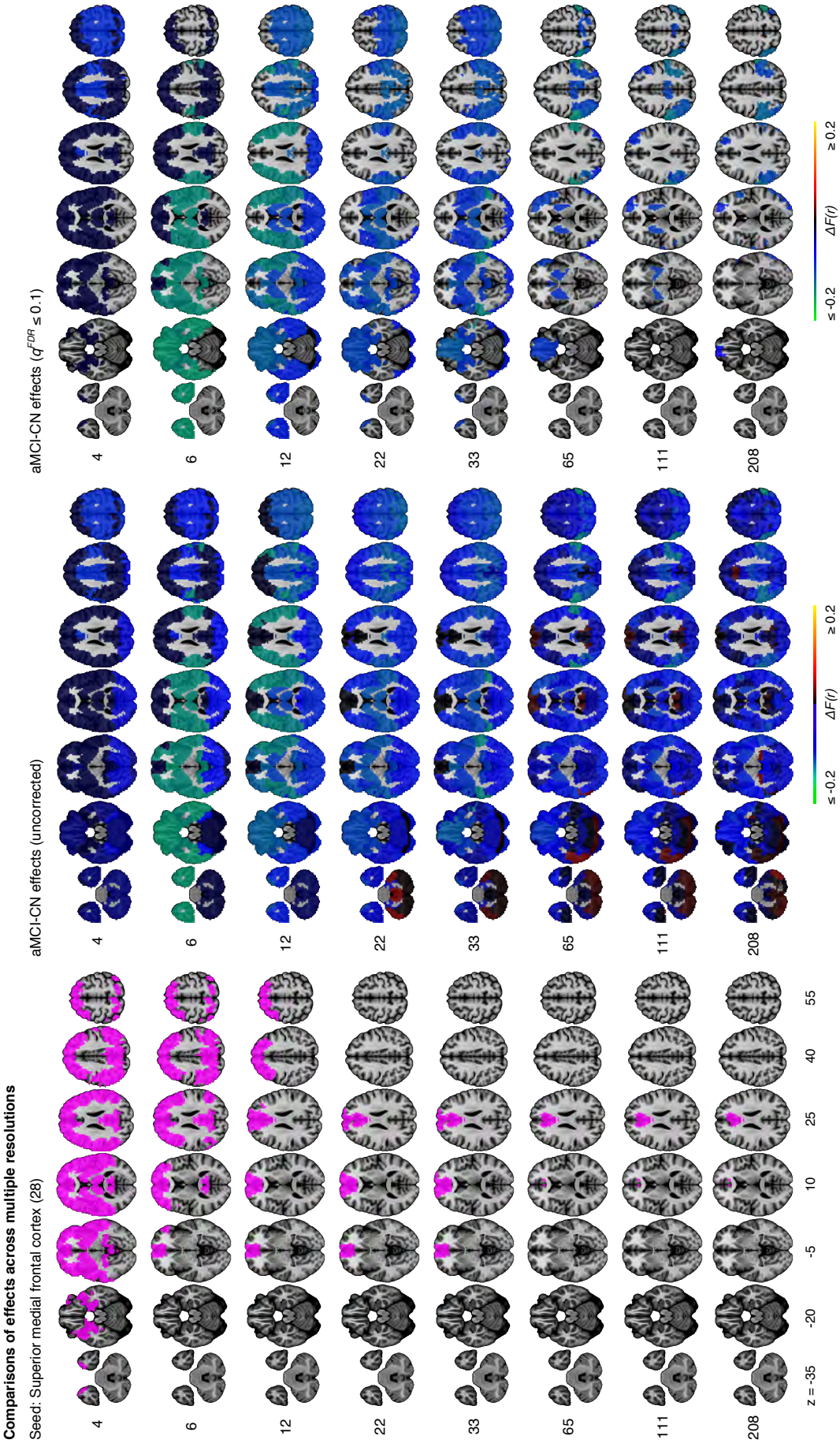

Figure 3.17: Comparisons of uncorrected and FDR-corrected effects related to aMCI across multiple resolutions selected by MSTEPS in the superior medial frontal cortex as the seed. The number in parentheses refers to the numerical ID of the seed in the 3D parcellation volume for 33 clusters, as listed in Supplementary Table 3.5.

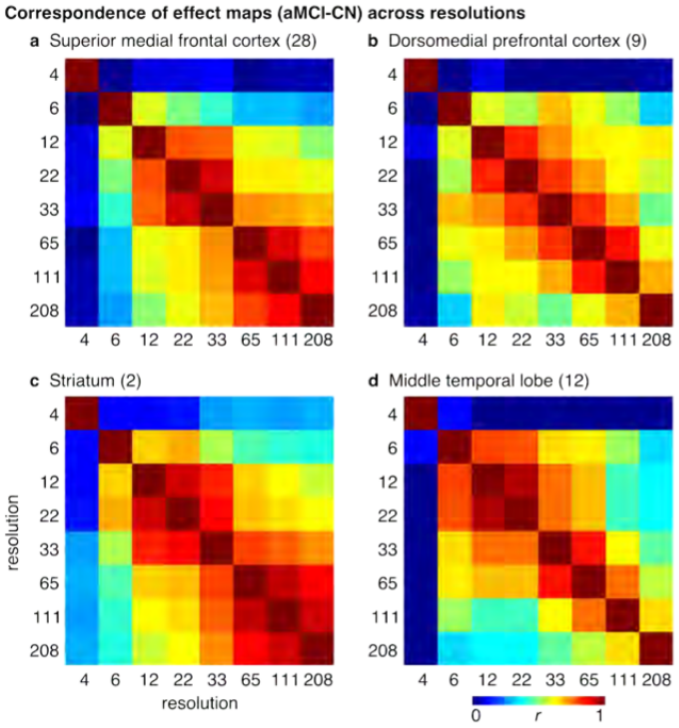

Figure 3.18: Correspondence of effect maps comparing aMCI against CN across resolutions selected by MSTEPS in four selected seeds. Correlation matrices show pairwise comparisons between the eight MSTEPS resolutions of the effect maps for the (A) superior medial frontal cortex, (B) dorsomedial prefrontal cortex, (C) striatum, and (D) middle temporal lobe. The numbers in parentheses refer to the numerical IDs of the seeds in the 3D parcellation volume for 33 clusters, as listed in Supplementary Table 3.5.

3.8.3 Supplementary tables

Table 3.4: Demographic information in all studies before quality control

		ADNI2	CRIUGMa	CRIUGMb	MNI	Combined
CN	**N**	59	20	18	15	112
	Mean age (s.d.)	74.1 (6.8)	71.7 (8.2)	69.9 (4.9)	67.0 (5.7)	72.0 (7.0)
	N male (%)	27 (46%)	7 (35%)	2 (11%)	7 (47%)	43 (38%)
	MMSE mean (range)	28.8 (24-30)	28.8 (27-30)	n/a	29.1 (27-30)	n/a
	MoCA mean (range)	n/a	27.6 (22-30)	28.5 (26-30)	n/a	n/a
aMCI	**N**	93	9	22	19	143
	Mean age (s.d.)	71.7 (7.8)	80.7 (6.2)	74.9 (6.9)	71.6 (8.1)	72.7 (7.7)
	N male (%)	47 (51%)	4 (44%)	13 (59%)	8 (42%)	72 (50%)
	MMSE mean (range)	28.1 (24-30)	26.3 (22-29)	n/a	25.9 (22-30)	n/a
	MoCA mean (range)	n/a	23.6 (20-29)	25.0 (16-29)	n/a	n/a

Table 3.5: Numerical IDs, labels, and volumes (mm^3) of parcels in the symmetric and asymmetric brain parcellations containing 33 clusters

ID	Label	Volume (Symmetric)	Volume (Asymmetric)
1	Thalamus	20547	20547
2	Striatum	47574	47574
3	Cerebellum_8_9_10	25461	25461
4	Angular/Parietal_Inf	34587	34074
5	Temporal_Pole	28242	28350
6	Cerebellum_3_4_5_6	42822	42768
7	Frontal_Inf_Tri	31860	31671
8	Cingulum_Post	38934	39204
9	Prefrontal_Dorsomedial	36693	35748
10	Hippocampus	57024	56997
11	Frontal_Sup	30969	31239
12	Temporal_Mid	32184	32157
13	Frontal_Mid/Sup	38043	36909
14	Supramarginal/Parietal_Inf	32670	31185
15	Cerebellum_Crus_Ant	48762	48843
16	Frontal_Inf_Oper	58239	57888
17	Postcentral	29538	29403
18	Precentral	44496	44334
19	Prefrontal_Ventrolateral	58158	58077
20	Precuneus	30051	30024
21	Fusiform/Parahippocampal	83997	82917
22	Prefrontal_Ventromedial	46035	47493
23	Frontal_Inf/Mid	57240	58266
24	Occipital_Mid/Inf	58671	61020
25	Temporal_Sup/Insula	65205	63666
26	Cuneus	60669	67230
27	Cerebellum_Crus_Post	76086	76086
28	Frontal_Sup_Medial	62424	62262
29	Calcarine/Lingual	57564	57483
30	Precentral/Supplementary_Motor	54135	54891
31	Pre/postcentral	70173	69606
32	Calcarine/Lingual/Cuneus	77328	77382
33	Occipital_Sup	80919	80190

Table 3.6: Rank of parcels based on their associated percentage of non-redundant connections that differ between aMCI and CN

Rank	Parcel	Additional percentage of connections	Cumulative percentage of connections
1	Superior medial frontal cortex (28)	13.44	13.44
2	Dorsomedial prefrontal cortex (9)	12.65	26.09
3	Striatum (2)	11.06	37.15
4	Middle temporal lobe (12)	9.89	47.04
5	Fusiform/Parahippocampal (21)	8.69	55.73
6	Angular/Inferior parietal (4)	7.12	62.85
7	Hippocampus (10)	6.32	69.17
8	Ventromedial prefrontal cortex (22)	6.32	75.49
9	Temporal pole (5)	4.75	80.24
10	Middle/Superior frontal cortex (13)	4.74	84.98
11	Superior temporal/Insula (25)	4.74	89.72
12	Inferior/Middle frontal cortex (23)	3.17	92.89
13	Prefrontal ventrolateral cortex (19)	2.37	95.26
14	Posterior cerebellar crus (27)	1.58	96.84
15	Calcarine/Lingual/Cuneus (32)	1.58	98.42
16	Anterior cerebellar crus (15)	0.79	99.21
17	Inferior frontal operculum (16)	0.79	100.00

Numbers in parentheses reference the number assigned to the cluster as per Supplementary Table 3.5

Chapter 4

Multiresolution functional brain parcellation in an elderly population with no or mild cognitive impairment

Angela Tam, Christian Dansereau, AmanPreet Badhwar, Pierre Orban, Sylvie Belleville, Howard Chertkow, Alain Dagher, Alexandru Hanganu, Oury Monchi, Pedro Rosa-Neto, Amir Shmuel, John Breitner, Pierre Bellec

4.1 Preface

This article acts as a companion paper to the study described in Chapter 3. The purpose of this paper was to describe and share functional parcellations derived from heterogeneous sets of resting-state fMRI data in older adults with or without mild cognitive impairment. We provide multiple parcellations at different spatial resolutions, spanning from large distributed networks to smaller regions of interest. These parcellations are available for download on figshare and Neurovault. These parcellations may be useful for researchers who are interested in resting-state functional connectivity in older populations. For example, a parcellation from this release could be used to define nodes in a graph analysis. This paper was published in Data in Brief in 2016[1].

4.2 Abstract

We present eight resolutions of group-level brain parcellations for clusters generated from resting-state functional magnetic resonance images for 99 cognitively normal

[1] http://dx.doi.org/10.1016/j.dib.2016.11.036

elderly persons and 129 patients with mild cognitive impairment, pooled from four independent datasets. This dataset was generated as part of the following study: Common Effects of Amnestic Mild Cognitive Impairment on Resting-State Connectivity Across Four Independent Studies (Tam et al., 2015). The brain parcellations have been registered to both symmetric and asymmetric MNI brain templates and generated using a method called bootstrap analysis of stable clusters (BASC) (Bellec et al., 2010). We present two variants of these parcellations. One variant contains bihemisphereic parcels (4, 6, 12, 22, 33, 65, 111, and 208 total parcels across eight resolutions). The second variant contains spatially connected regions of interest (ROIs) that span only one hemisphere (10, 17, 30, 51, 77, 199, and 322 total ROIs across eight resolutions). We also present maps illustrating functional connectivity differences between patients and controls for four regions of interest (striatum, dorsal prefrontal cortex, middle temporal lobe, and medial frontal cortex). The brain parcels and associated statistical maps have been publicly released as 3D volumes, available in .mnc and .nii file formats on figshare and on Neurovault. Finally, the code used to generate this dataset is available on Github.

4.3 Data

This data release contains group brain parcellations at multiple resolutions (4, 6, 12, 22, 33, 65, 111, and 208 parcels) generated from resting-state functional magnetic resonance images for 99 cognitively normal elderly persons and 129 patients with mild cognitive impairment. This work also includes parcellations that contain regions-of-interest (ROIs) that are spatially connected and span only one hemisphere at 8 resolutions (10, 17, 30, 51, 77, 137, 199, and 322 total ROIs). Labels based on typical resting-state networks, and their decomposition into subnetworks or regions, are proposed for all brain parcels. This release also includes unthresholded maps of connectivity differences (t-maps) between patients and controls for four seeds/regions of interest (striatum, dorsal prefrontal cortex, middle temporal lobe, and medial frontal cortex).

4.4 Experimental design, materials and methods

4.4.1 Participants

We pooled resting-state functional magnetic resonance imaging (fMRI) data from four independent studies: the Alzheimer's Disease Neuroimaging Initiative 2 (ADNI2) sample, two samples from the Centre de recherche de l'institut universitaire de

gériatrie de Montréal (CRIUGMa and CRIUGMb), and a sample from the Montreal Neurological Institute (MNI) (3). All participants gave their written informed consent to engage in these studies, which were approved by the research ethics board of the respective institutions, and included consent for data sharing with collaborators as well as secondary analysis. Ethical approval was also obtained at the site of secondary analysis (CRIUGM). The ADNI2 data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership representing efforts of co-investigators from numerous academic institutions and private corporations. ADNI was followed by ADNI-GO and ADNI-2 that included newer techniques. Subjects included in this study were recruited by ADNI-2 from all 13 sites that acquired resting-state fMRI on Philips scanners across North America. For up-to-date information, see www.adni-info.org. The final combined sample included 112 cognitively normal elderly subjects (CN) and 143 patients with mild cognitive impairment (MCI). In the CN group, the mean age was 72.0 (s.d. 7.0) years, and 38.4% were men. Mean age of the MCI subjects was 72.7 (s.d. 7.7) years, and 50.3% were men. For more information about recruitment or participant characteristics, please refer to Tam et al. (2015).

4.4.2 Imaging data acquisition

All resting-state fMRI and structural scans were acquired on Philips and Siemens 3T scanners. For more detailed information on the imaging parameters, please refer to Tam et al. (2015).

4.4.3 Computational environment

All experiments were performed using the NeuroImaging Analysis Kit (NIAK)[2] (Bellec et al., 2011) version 0.12.18, under CentOS version 6.3 with Octave[3] version 3.8.1 and the Minc toolkit[4] version 0.3.18. Analyses were executed in parallel on the "Guillimin" supercomputer[5], using the pipeline system for Octave and Matlab (Bellec et al., 2012), version 1.0.2. The scripts used for processing can be found on Github[6].

[2] http://simexp.github.io/niak/
[3] https://www.gnu.org/software/octave/
[4] http://www.bic.mni.mcgill.ca/ServicesSoftware/ServicesSoftwareMincToolKit
[5] http://www.calculquebec.ca/en/resources/compute-servers/guillimin
[6] https://github.com/SIMEXP/mcinet

4.4.4 Pre-processing

Each fMRI dataset underwent preprocessing as described in (Tam et al., 2015). A more detailed description of the pipeline can also be found on the NIAK website[7] and Github[8].

4.4.5 Parcellation of the brain into functional clusters

After pre-processing, we generated functional brain atlases at eight resolutions, containing 4, 6, 12, 22, 33, 65, 111 and 208 total parcels, as described in (Tam et al., 2015). These eight resolutions of brain parcellations (Figure 4.1), registered to both symmetric and asymmetric MNI templates, have been released on figshare[9] and Neurovault[10]. These eight resolutions were further processed to generate eight parcellations that contain ROIs that are spatially connected and span only one hemisphere (for an example, see Figure 4.2). These latter parcellations contain 10, 17, 30, 51, 77, 137, 199, and 322 total ROIs.

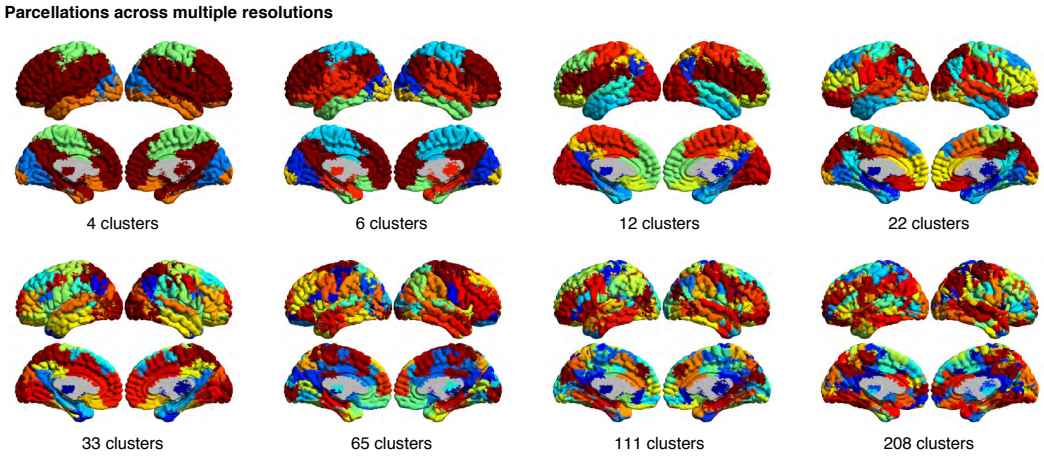

Figure 4.1: Functional parcellations across resolutions (or number of clusters).

We have provided labels for each parcel at every resolution, except for resolutions 4 and 6 due to the merging of networks at those low resolutions. At resolution 4, we observed the sensory-motor network, visual network, a network that resembles the endogenous network (6) and a network that merges the cerebellum and the mesolimbic network together. At resolution 6, we observed the visual network, cerebellum, mesolimbic network, sensory-motor network, a network that merges the deep gray matter nuclei with the frontoparietal network, and a network that merges the default mode network with the posterior attention network. For resolution 12,

[7] http://niak.simexp-lab.org/pipe_preprocessing.html
[8] https://github.com/SIMEXP/mcinet/tree/master/preprocess
[9] http://dx.doi.org/10.6084/m9.figshare.1480461
[10] http://neurovault.org/collections/1003/

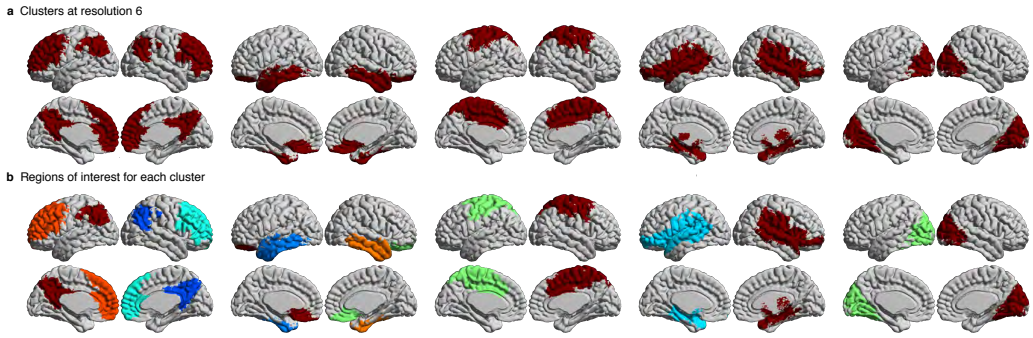

Figure 4.2: Clusters at resolution 6 (cerebellum not shown) and their respective regions-of-interest. Note how each cluster in (A) is bihemispheric prior to breaking down into multiple spatially constrained regions-of-interest in (B).

we manually labeled each parcel (deep gray matter nuclei (DGMN), posterior default mode network (pDMN), medial temporal lobe (mTL), ventral temporal lobe (vTL), dorsal temporal lobe (dTL), anterior default mode network (aDMN), orbitofrontal cortex (OFC), posterior attention (pATT), cerebellum (CER), sensory-motor (SM), visual (VIS), and frontoparietal network (FPN)). Then, we decomposed the networks at resolution 12 into smaller subclusters at all higher resolutions (for an example, see Figure 4.3). Each parcel at higher resolutions was labeled in reference to the parcels at resolution 12, with the following convention: (resolution)_(parcel label)_(#); for example, at resolution (R) 22, the anterior default mode splits into two clusters, which were named "R22_aDMN_1" and "R22_aDMN_2".

4.4.6 Derivation of functional connectomes

Between and within-clusters connectivity was measured as described in Tam et al., 2015.

4.4.7 Statistical testing

To test for differences between aMCI and CN at a resolution of 33 clusters, we used a general linear model (GLM) for each connection between two parcels (Bellec et al., 2015). Specific details of the GLM can be found in (Tam et al., 2015). From this analysis, we present uncorrected t-maps illustrating functional connectivity differences between patients and controls for four seeds/regions of interest (striatum, dorsal prefrontal cortex, middle temporal lobe, and medial frontal cortex) (Figure 4.4). These maps have been released on figshare and Neurovault. These four seeds were chosen for further analyses because, together, they were associated with 47% of all significant group differences across all brain regions. Briefly, we found that MCI patients exhibited reduced connectivity between default mode network nodes and

between areas of the cortico-striatal-thalamic loop. For a more in-depth presentation and discussion of results, please refer to Tam et al. (2015).

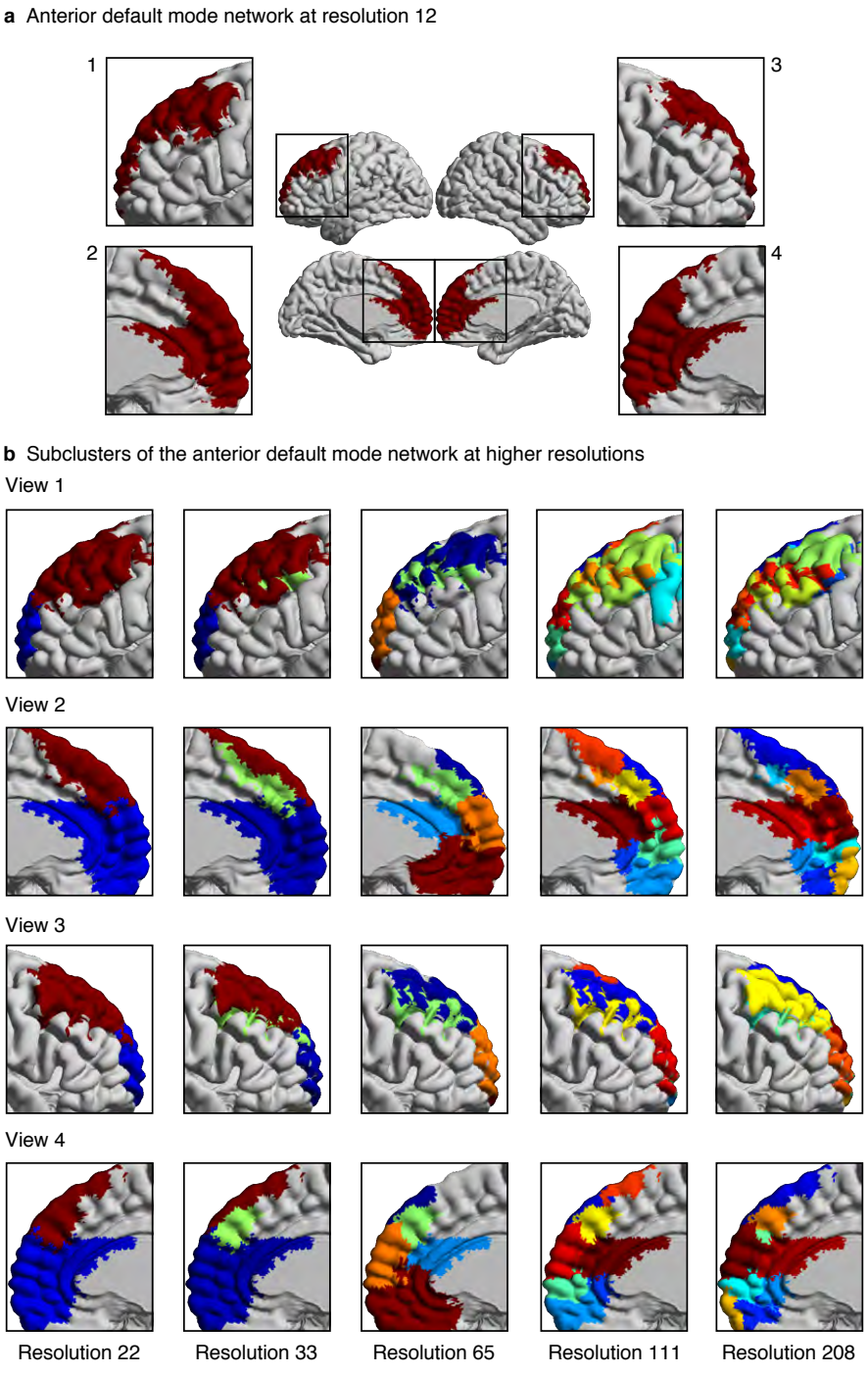

Figure 4.3: The decomposition of the anterior default mode network into smaller subclusters at higher resolutions in four different views. Resolution 12 was used as a reference for the labeling of subnetworks at higher resolutions.

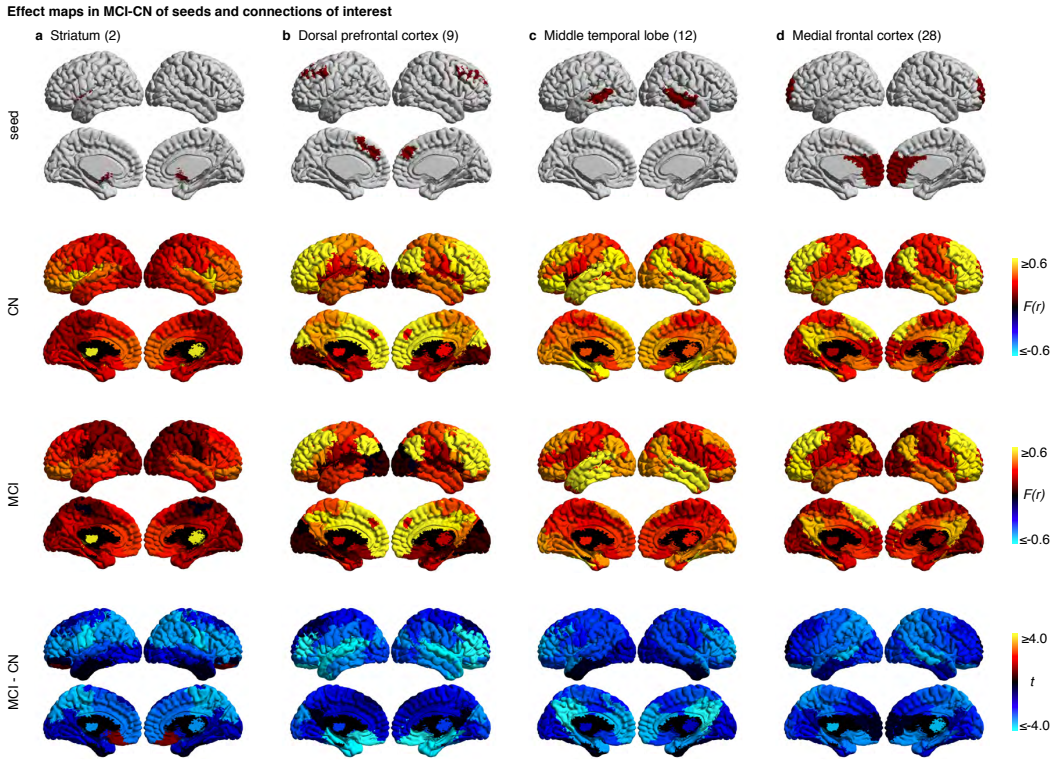

Figure 4.4: Maps for a selection of four seeds that show effects related to MCI at resolution 33. These effect maps reveal the spatial distribution of the differences in functional connectivity for (A) striatum, (B) dorsal prefrontal cortex, (C) middle temporal lobe, and (D) the medial frontal cortex. For each panel, the top line maps the spatial location of the seed region in red, the second and third lines show the connectivity (Fisher-transformed correlation values ($F(r)$) between the designated seed region and the rest of the brain in CN and MCI respectively, and the fourth line shows a difference map between MCI and CN (t-test). The numbers in parentheses refer to the numerical IDs of the clusters in the 3D parcellation volume at resolution 33.

4.5 Acknowledgements

We thank Sven Joubert, Isabelle Rouleau, and Sophie Benoit for their contribution in collection of the CRIUGMa dataset. We thank Seqian Wang for his help in curating the ADNI dataset. We thank Tom Nichols for his comments and suggestions. Collection of the CRIUGMb dataset was supported by the Alzheimer's Society of Canada. Collection of the MNI dataset was supported by Industry Canada/Montreal Neurological Institute Centre of excellence in commercialization and research. This work was supported by CIHR (133359) and NSERC (436141-2013). The computational resources used to perform the data analysis were provided by Compute Canada and CLUMEQ, which is funded in part by NSERC (MRS), FQRNT, and McGill University.

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 DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing 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 coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Chapter 5

A signature of cognitive deficits and brain atrophy that is highly predictive of progression to Alzheimer's dementia

Angela Tam, Christian Dansereau, Yasser Iturria-Medina, Sebastian Urchs, Pierre Orban, John Breitner, Pierre Bellec

5.1 Preface

In Chapter 3, we found a large amount of variation in functional connectivity across different samples of MCI patients and cognitively normal subjects. A number of studies had also reported different subtypes of brain atrophy patterns within AD patients. As brain structure and function are often related, it seemed plausible that the heterogeneity across samples with respect to functional connectivity could be due to variability similar to that found in atrophy. To explore this, we conducted a study that examined subtypes of functional connectivity and whether certain subtypes may inform an individual's risk of developing AD from a preclinical stage (see Orban et al. (2017) in Appendix B). We next conducted a second study combining functional and structural subtypes with a novel machine learning algorithm to generate a predictive model to identify MCI subjects who would eventually progress to AD dementia with high specificity and positive predictive value (see Dansereau et al. (2017) in Appendix C). This current study is a follow-up to these last two works. We aimed to develop a multimodal signature, based on cognition and subtypes of grey matter atrophy, that is highly predictive of progression to AD dementia in a subgroup of MCI patients. We

chose to use tools that were regularly used in clinical settings (i.e. structural MRI and neuropsychological assessments) in order to evaluate the complementarity of features derived from cognition and atrophy patterns. We also applied the machine learning algorithm from Dansereau et al. (2017) to validate it on several larger datasets. This paper will be submitted for publication at a peer-reviewed journal.

5.2 Abstract

Individuals suffering from mild cognitive impairment (MCI) have an increased risk of developing Alzheimer's disease (AD) dementia; yet only a fraction of them do. We explore here whether a high risk subgroup can be identified using a signature, comprised of spatial patterns of grey matter atrophy and cognitive deficits. We applied machine learning algorithms to identify such a signature that it is commonly seen in ADNI1 patients with AD dementia but absent in cognitively normal subjects. We validated the signature in an independent cohort of AD and control subjects from ADNI2. We then applied the signature on two separate cohorts of MCI subjects, from ADNI1 and ADNI2, with the hypothesis that it would distinguish individuals who progressed to dementia from those who remained cognitively stable over three years. In ADNI1, the model predicted progression to dementia in MCI patients with 80.4% positive predictive value, adjusted for a "typical" baseline rate of 33%, 95.6% specificity and 55.1% sensitivity. These results were replicated in ADNI2, where the model reached 87.8% adjusted positive predictive value, 96.7% specificity, and 47.3% sensitivity. This signature was optimized for high specificity, which resulted in higher positive predictive values compared to previous work, at the expense of sensitivity. The signature suggests that cognitive deficits across multiple domains combined with more atrophy in temporal, parietal and occipital lobe regions are factors that are highly predictive of future progression. Our signature, which used non-invasive and widely available markers, could be used for subject selection in clinical trials or identification of high-risk individuals for earlier intervention.

5.3 Introduction

Alzheimer's disease (AD) is an age-related neurodegenerative disorder and a leading cause of dementia worldwide. AD is marked by the abnormal accumulation of amyloid β (Aβ) and hyperphosphorylated tau proteins in the brain, which leads to widespread neuronal damage. These neurobiological events may develop years, or even decades, before the emergence of mild cognitive impairment (MCI), considered a prodromal stage of AD. MCI is a clinical condition not specific of AD;

many patients will remain cognitively stable over time or even revert to normal cognition. A diagnosis of MCI is still a risk factor for AD dementia, with up to 36% of patients developing dementia within two years (Ward et al., 2013). Given the long preclinical and prodromal phases in AD, there is interest in developing treatments that could prevent further decline, which requires biomarkers predictive of progression to dementia from preclinical stages. Identifying progressor MCI to AD dementia with enough specificity has been a challenge for clinical trials, where the inclusion criteria for MCI subjects have had low to moderate positive predictive value (Visser, Scheltens, and Verhey, 2005). This lack of prognostic power could be due to underlying heterogeneity among individuals. Dong et al. (2017) for example showed four distinct patterns of brain atrophy, two of which were associated with higher-than-expected rates of progression to dementia. Yet, one subtype was much more at risk than the other. The implications for prognosis are profound: only a subset of individuals will have clinical trajectories that can be predicted very reliably. In this work, we propose to apply machine learning tools to identify subgroups at very high risk of progression to AD dementia, using commonly available measures, i.e. structural magnetic resonance imaging (MRI) and cognitive tests.

There is a growing field of research dedicated to developing machine learning algorithms to automatically detect progression to AD dementia in patients with MCI. Cognitive tests measuring episodic memory and language have been identified as features with high predictive accuracy (Belleville et al., 2017), with specificity and sensitivity values of at least 70%. It has been suggested that the combination of several cognitive domains, notably memory, language, and executive function tests, yields the best performance and balance between specificity and sensitivity (Belleville et al., 2017). With structural MRI, it is well established that significant tissue loss occurs in the medial temporal lobes in AD patients (Jack et al., 1997). Medial temporal lobe atrophy may be an early marker of AD as it can be seen also in MCI patients relative to cognitively normal individuals (McDonald et al., 2009). As the disease progresses and clinical symptoms become more severe, atrophy spreads to the parietal and frontal lobes, while leaving the sensorimotor and visual cortices relatively intact (McDonald et al., 2009). Grey matter volumes (GMV) of regions of interest, typically defined by an atlas, have been used as features for the classification of progressor MCI (pMCI) and stable MCI (sMCI) (Misra, Fan, and Davatzikos, 2009; Davatzikos et al., 2011). Cortical thickness of brain regions have similarly been used as features to classify sMCI and pMCI (Eskildsen et al., 2013; Querbes et al., 2009; Wee et al., 2013). Models using structural features alone have achieved variable performance, ranging from 38-84% specificity and 63-95% sensitivity. Models that have combined structural MRI markers with cognitive measurements have

outperformed models with either structural MRI or cognition alone at predicting progression to AD dementia (Devanand et al., 2007; Korolev et al., 2016; Moradi et al., 2015). For models combining structural MRI and cognition, in terms of accuracy, the best has been reported at 79% (76% specificity, 83% sensitivity) (Korolev et al., 2016). Overall, the state-of-the-art, which used Aβ positron emission tomography (PET) scans, has so far achieved 82% accuracy (87% specificity, 71% sensitivity) (Mathotaarachchi et al., 2017). Note that accuracy is defined as the proportion of subjects that were correctly identified, either as positives or negatives. Positive predictive value, on the other hand, is defined as the proportion of true positives out of all subjects that were classified as positive (including the false negatives). A model could have high accuracy yet moderate positive predictive value if the number of true positives is low relative to the number of true negatives. In the case of predicting incipient AD, relatively few MCI subjects progress to dementia (up to 36% within two years (Ward et al., 2013)). Despite promising accuracy, the positive predictive values of models predicting AD progression remain moderate, ranging from 50 to 75% across the literature (Dansereau et al., 2017). This implies that up to half of subjects who were identified as progressors by published algorithms did not in fact progress to dementia. Furthermore, predictive power decreases as the time to conversion increases (Eskildsen et al., 2013), so there is even more room for improvement for long-term prognosis (beyond 2 years).

AD dementia prognosis may be difficult due to the inherent heterogeneity present across the population. While brain amyloidosis is a defining feature of AD, it is clearly insufficient at causing clinical impairment since 10-44% of cognitively normal subjects aged 50 to 90 years are amyloid-positive, based on a meta-analysis of 55 studies with PET and cerebrospinal fluid (Jansen et al., 2015). Furthermore, a meta-analysis of 29 cohorts revealed that, on average, 88% of clinically diagnosed AD patients have amyloid-negative PET scans (Ossenkoppele et al., 2015). Similar prevalence rates have been corroborated by post-mortem neuropathological assessments (Beach et al., 2012; Bennett et al., 2006), revealing that clinical diagnoses often do not match with pathology. Beach et al. (2012) also found that guidelines for a neuropathological diagnosis of AD were too restrictive, rendering 33% of cases unclassifiable. It is apparent that we need to better characterize heterogeneity within populations if we are to reach the precision levels required for individualized medicine.

To address some of this heterogeneity, several studies have explored subtypes of atrophy in AD, where homogeneous subgroups of AD patients show different patterns of gray matter loss relative to healthy controls. For example, a number of works have reproduced a subtype characterized by primarily medial temporal lobe atrophy, a second subtype characterized by parietal, occipital, and lateral temporal

lobe atrophy, and a third subtype with diffuse atrophy (Noh et al., 2014; Hwang et al., 2016; Park et al., 2017; Varol et al., 2017). The subgroup with parietal, occipital, and lateral temporal lobe atrophy tended to be younger and had the worst cognition (Noh et al., 2014; Park et al., 2017), while the medial temporal lobe atrophy subtype contained a greater proportion of carriers of the APOE4 allele (Varol et al., 2017), a genetic risk factor for sporadic late-onset AD, and the diffuse atrophy subtype had the least Aβ and tau burden (Varol et al., 2017). Heterogeneity is also present in the symptomatic expression of MCI patients. Often, MCI patients are categorized as amnestic MCI, subjects who have memory impairment, non-amnestic MCI, those who have impairments in non-memory domains, and multi-domain MCI, subjects with concurrent impairments in multiple kinds of cognition (Petersen, 2003). This would result in a patient being labeled as one of four MCI subtypes: 1) amnestic single domain MCI, 2) non-amnestic single domain MCI, 3) amnestic multidomain MCI, 4) non-amnestic multidomain MCI (Petersen et al., 2014). Amnestic MCI subjects are more likely to be diagnosed with dementia than non-amnestic MCI (Aggarwal et al., 2005; Busse et al., 2006), but the group at highest risk of conversion is multi-domain MCI (Tabert et al., 2006). Mitchell et al. (2009) reported 59% of multi-domain MCI subjects progressed after two years of follow-up, while only 18% of amnestic single domain MCI progressed, and 70% of non-amnestic single domain MCI actually improved, which illustrates how outcomes can vary substantially between subgroups. Certain subtypes of brain patterns may predispose individuals to worse cognitive trajectories. Subjects with widespread brain atrophy, and especially those who also express greater loss in the temporal lobes, have worse cognitive decline compared to other subjects with more localized atrophy (Dong et al., 2017).

While brain structure and cognition appear to offer complementary information for prognosis of dementia, the complementarity of those modalities to identify high-risk subtypes is unclear. It is possible that a predictive signature based on atrophy may identify a subgroup of subjects that is separate from those labeled by a signature based on cognition, for example. Then despite the many works that describe subtypes of brain atrophy in AD populations (Noh et al., 2014; Hwang et al., 2016; Park et al., 2017; Varol et al., 2017; Zhang et al., 2016; Dong et al., 2016), little has been done to leverage the heterogeneity to predict incipient dementia from a subgroup of MCI patients with high positive predictive value. We have previously shown that the combined expression of certain subtypes of cortical thinning patterns, grey matter volumes, and resting-state network functional connectivity was highly predictive of future AD progression in a subgroup of MCI subjects with 90% positive predictive value (Dansereau et al., 2017). While the results from Dansereau et al. (2017) are

promising, they require further validation since the sample size was small as resting-state functional MRI data is relatively uncommon compared to other measures.

The main aim of this work was to develop a multimodal signature, based on cognition and subtypes of grey matter atrophy, that is highly predictive of progression to AD dementia in a subgroup of MCI patients. We wanted to evaluate the complementarity of features derived from cognition and atrophy patterns. We applied a cluster analysis on structural magnetic resonance images to identify subtypes of brain atrophy in a sample containing both patients with AD dementia and cognitively normal (CN) individuals. We then used a two-step machine learning algorithm (Dansereau et al., 2017) to train a model to identify three signatures that were highly predictive of AD dementia: 1) an anatomical signature, 2) a cognitive signature, 3) a multimodal anatomical and cognitive signature. After identifying MCI patients carrying these signatures, we examined cognitive decline, Aβ and tau burden, and progression to dementia in these individuals to explore whether a highly predictive signature represented a prodromal stage of AD. We analysed whether these three signatures identified separate subgroups of subjects and how they performed against each other in terms of positive predictive value, specificity, and sensitivity at identifying progressors.

5.4 Methods

5.4.1 Data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org. We took baseline T1-weighted MRI scans from the ADNI1 (228 CN, 397 MCI, 192 AD) and ADNI2 (218 CN, 354 MCI, 103 AD) studies. For a detailed description of MRI acquisition details, see http://adni.loni.usc.edu/methods/documents/mri-protocols/. All subjects gave informed consent to participate in these studies, which were approved by the research ethics committees of the institutions involved in data acquisition. Consent was obtained for data sharing and secondary analysis, the latter being approved by the ethics committee at the CRIUGM. For the MCI groups, each individual must have

Table 5.1: Demographic information for post-QC subjects in ADNI1 and ADNI2.

ADNI1	**CN**	**sMCI**	**pMCI**	**AD**
N	205	89	155	172
Age ± SD	76.1 ± 5.0	73.9 ± 7.6	74.2 ± 7.1	75.3 ± 7.5
Female %	51.7	40.4	41.9	50.0
APOE4+ %	27.8	37.1	68.4	66.3
ADAS13 ± SD	9.5 ± 4.3	14.4 ± 5.5	21.3 ± 5.2	28.7 ± 7.1
MMSE ± SD	29.1 ± 1.0	27.7 ± 1.7	26.7 ± 1.7	23.4 ± 2.0
ADNI2	**CN**	**sMCI**	**pMCI**	**AD**
N	188	182	55	90
Age ± SD	72.8 ± 6.1	70.8 ± 7.3	72.1 ± 7.1	74.6 ± 7.9
Female %	54.0	47.3	49.1	45.6
APOE4+ %	29.4	36.3	65.4	70.4
ADAS13 ± SD	9.1 ± 4.2	11.8 ± 5.3	21.4 ± 6.5	31.9 ± 8.7
MMSE ± SD	29.1 ± 1.1	28.4 ± 1.5	27.3 ± 1.9	23.1 ± 2.3

ADAS13 = Alzheimer's Disease Assessment Scale - Cognitive subscale (13 items); MMSE = Mini Mental State Examination

had at least 36 months of follow-up for inclusion in our analysis. We also further stratified the MCI groups into stable (sMCI), who never received any change in their diagnosis, and progressors (pMCI), who received a diagnosis of AD dementia within 36 months of follow-up. pMCI who progressed to AD dementia after 36 months were excluded. After applying these inclusion/exclusion criteria, we were left with 280 and 268 eligible MCI subjects in ADNI1 and ADNI2 respectively.

Structural features from voxel-based morphometry

Images were processed and analyzed with the NeuroImaging Analysis Kit (NIAK) version 0.18.1 (https://hub.docker.com/r/simexp/niak-boss/), the MINC toolkit (http://bic-mni.github.io/) version 0.3.18, and SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) under CentOS with Octave (http://gnu.octave.org) version 4.0.2. Preprocessing of MRI data was executed in parallel on the Cedar supercomputer (https://docs.computecanada.ca/wiki/Cedar), using the pipeline system for Octave and Matlab - PSOM (Bellec et al., 2012). Each T1 image was linearly co-registered to the Montreal Neurological Institute (MNI) ICBM152 stereotaxic symmetric template (Fonov et al., 2011), using the CIVET pipeline (Ad-Dab'bagh et al., 2006), and then re-oriented to the AC-PC line. Each image was segmented into grey matter, white matter, and CSF probabilistic maps. The DARTEL toolbox (Ashburner, 2007) was used to normalize the grey matter segmentations to a predefined grey matter template in MNI152 space. Each map was modulated to preserve the total amount of signal and smoothed with with a 8 mm isotropic Gaussian blurring kernel. After quality control of the normalized grey matter segmentations, we were left with 621 subjects in ADNI1 (out of 700, 88.7% success rate) and 515 subjects in ADNI2 (out of 589, 87.4% success rate). See Table 5.1 for demographic information of post-QC subjects.

We extracted subtypes to characterize variability of grey matter distribution with the CN and AD samples from ADNI1. In order to reduce the impact of factors of no interest that may have influenced the clustering procedure, we regressed out age, sex, mean grey matter volume (GMV), and total intracranial volume (TIV), using a mass univariate linear regression model at each voxel. We then derived a spatial Pearson's correlation coefficient between all pairs of individual maps after confound regression. This defined a subject x subject (377 x 377) similarity matrix which was entered into a Ward hierarchical clustering procedure (Figure 5.1a). Based on visual inspection of the similarity matrix, we identified 7 subgroups (Figure 5.1b). Each subtype was defined as the average map of each subgroup. We then computed spatial correlations between each individual's map and each subtype, referred to as weights (Figure 5.1a). These weights were used as features for the predictive models involving VBM in the rest of this work. As in our previous works (Orban et al., 2017; Dansereau et al., 2017), we chose to use weights, which can be interpreted as continuous measures for subtype affinity, over discrete subtype membership because the latter is less informative as most individuals express similarity to multiple subtypes (Zhang et al., 2016). Note that although we chose to present our findings with 7 subtypes, we examined how the number of subtype may impact our subsequent predictions. There was no significant difference in model performance when we changed the number of subtypes (see Table 5.3 in supplementary material).

Cognitive features

We took baseline neuropsychological scores for each subject from several cognitive domains: memory from the composite score ADNI-MEM (Crane et al., 2012), executive function from the composite score ADNI-EF (Gibbons et al., 2012), language from the Boston Naming Test (BNT), visuospatial from the clock drawing test, and global cognition from the Alzheimer's Disease Assessment Scale-Cognitive (ADAS13). We chose measures that span multiple cognitive domains as it has been suggested that the use of a combination of neuropsychological measures is likely the best approach to predicting incipient dementia (Belleville et al., 2017). These scores were included as features for the predictive models involving cognition.

Prediction of easy AD dementia cases in ADNI1

We trained a support vector machine (SVM) model with a linear kernel, as implemented by Scikit-learn (Pedregosa et al., 2011) version 0.18 to classify AD vs CN from ADNI1 to get a baseline prediction accuracy. A tenfold cross-validation loop was used to estimate the performance of the trained model. Classes were balanced

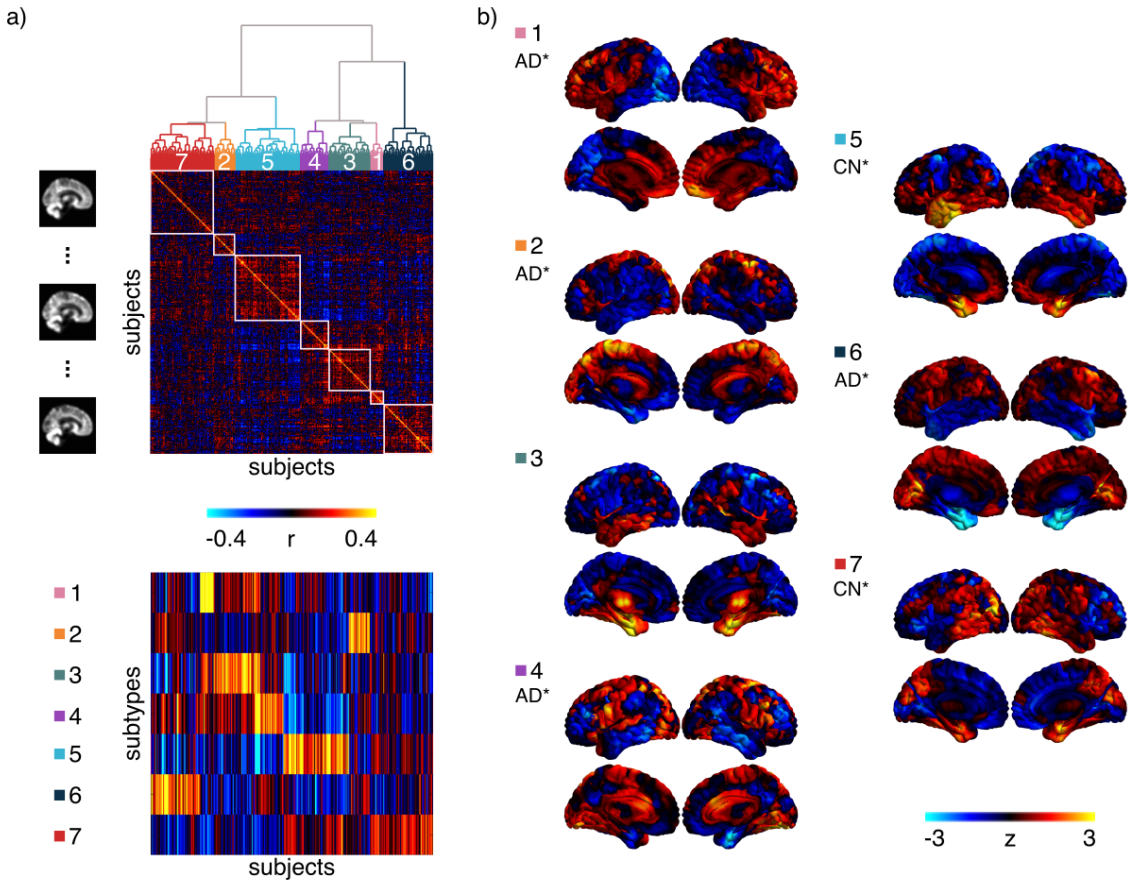

Figure 5.1: Subtyping procedure and resulting subtypes. a) A hierarchical clustering procedure identified 7 subtypes, or subgroups, of individuals with similar patterns of grey matter topography within the ADNI1 cohort of CN and AD subjects (top). A measure of spatial similarity, called subtype weight, between a single individual's grey matter volume map and the average of a given subtype was calculated for all individuals and all subtypes (bottom). b) Maps of the 7 subtypes showing the distribution of grey matter across all voxels relative to the average. CN* and AD* denote significant associations between the subtype weights and diagnoses of cognitively normal (CN) or Alzheimer's dementia (AD) respectively.

inversely proportional to class frequencies in the input data for the training. A nested cross-validation loop (stratified shuffle split with 50 splits and 20% test size) was used for the grid search of the hyperparameter C (grid was 10^{-2} to 10^{1} with 15 equal steps). We randomly selected subsamples of the dataset (retaining 50% of participants in each subsample) to replicate the SVM training 500 times. For each 50% subsample, a separate SVM model was trained to predict AD or CN in ADNI1. Predictions were made on the remaining 50% of the sample that was not used for training. For each subject, we calculated a hit probability defined as the frequency of correct classification across all SVM replications in which the test set contained that subject. Easy AD cases were defined as individuals with 100% hit probabilities with the AD label. Next, we trained a logistic regression classifier (Fan et al., 2008), with L1 regularization on the coefficients, to predict the easy AD cases. A stratified shuffle split (500 splits, 50% test size) was used to estimate the performance of the model for the grid search of the hyperparameter C (grid was 10^{-2} to 10^{1} with 15 equal steps). See (Dansereau et al., 2017) for more information about this two-step prediction. We used the whole CN and AD sample from ADNI1 to obtain three highly predictive signatures (HPS), one using VBM subtypes (VBM only), one using cognitive features (COG only), and one using the combination of VBM and cognitive features (VCOG). In all three signatures, age, sex, mean GMV, and TIV were also included as features.

5.4.2 Prediction of progression to AD dementia from the MCI stage in ADNI1

The logistic regression trained on AD vs CN was used to identify MCI patients who have a HPS of AD dementia in ADNI1. We re-trained our models on AD vs CN after optimizing our hyperparameters (resampling size and resampling ratio) in order to maximize specificity and positive predictive value while keeping a minimum of 30% sensitivity for our classification of sMCI (n=89) vs pMCI (n=155) in ADNI1. This was done for all three signatures. In brief, we used the AD and CN sample from ADNI1 as a training set, the MCI subjects from ADNI1 as a validation set, and ADNI2 served as our test set.

5.4.3 Statistical test of differences in model performance

We used Monte-Carlo simulations to generate confidence intervals on the performance (i.e. positive predictive value, specificity and sensitivity) of both base SVM and HPS models for their predictions of AD vs CN and pMCI vs sMCI. Taking the observed sensitivity and specificity, and using similar sample sizes to our experiment,

we replicated the number of true and false positive detection 100000 times using independent Bernoulli variables, and derived replications of positive predictive value, specificity and sensitivity. By comparing these replications to the sensitivity, specificity and positive predictive value observed in both models, we estimated a p-value for differences in model performance (Phipson and Smyth, 2010). A p-value smaller than 0.05 was interpreted as evidence of a significant difference in performance, and 0.001 as strong evidence. We also used this approach to compare the performance of the combined features (VCOG) to the models containing VBM features or cognitive features only. Note that, based on our hypotheses regarding the behaviour of the HPS model, the tests were one-sided for increased specificity and positive predictive value, and one-sided for decreased sensitivity.

5.4.4 Statistical tests of association of progression, AD biomarkers, and risk factors in HPS+ MCI subjects

Based on the classifications resulting from the base SVM and HPS models, we separated the MCI subjects into three different groups: 1) HPS+, subjects who were selected by the HPS model as hits, 2) Non-HPS+, subjects who were selected by the base SVM model as hits but were not selected by the HPS model, and 3) Negative, subjects who were not selected as hits by either algorithm.

We tested statistically if the HPS+ subgroup was enriched for progression to dementia, APOE4 carriers, females, and subjects who were positive for Aβ and tau pathology. Positivity of AD pathology was determined by CSF measurements of Aβ 1-42 peptide and total tau with cut-off values of less than 192 pg/mL and greater than 93 pg/mL respectively (Shaw et al., 2009). We implemented Monte-Carlo simulations, where we selected 100000 random subgroups out of the original MCI sample. By comparing the proportion of progressors, APOE4 carriers, females, Aβ-positive, and tau-positive subjects in these null replications to the actual observed values in the HPS subgroup, we estimated a p-value (Phipson and Smyth, 2010) (one sided for increase). A p-value smaller than 0.05 was interpreted as evidence of a significant enrichment, and 0.001 as strong evidence.

One-way ANOVAs were used to evaluate differences between the HPS groupings with respect to age. Post-hoc Tukey's HSD tests were done to assess pairwise differences among the three classes (HPS+, Non-HPS+, Negative). These tests were implemented in Python with the SciPy library (Jones, Oliphant, and Peterson, 2001) version 0.19.1 and StatsModels library (Seabold and Perktold, 2010) version 0.8.0.

To explore the impact of HPS grouping on cognitive trajectories, linear mixed effects models were performed to evaluate the main effects of and interactions

between the HPS groups and time on ADAS13 scores up to 36 months of follow-up. The models were first fit with a random effect of participant and then were fit with random slopes (time | participant) if ANOVAs comparing the likelihood ratio suggested a significant improvement in model fit. All tests were performed separately on the ADNI1 and ADNI2 datasets. These tests were implemented in R version 3.3.2 with the library nlme version 3.1.128 (Pinheiro et al., 2018).

5.4.5 Public code and data

The code used in this experiment is available on a Github repository: https://github.com/SIMEXP/vcog_hps_ad.

5.5 Results

5.5.1 Prediction of AD dementia vs cognitively normal individuals

A linear SVM model trained using the VCOG features achieved 94.5% positive predictive value (95.6% specificity, 93.9% sensitivity) when classifying AD vs CN in ADNI1. Such high performance was expected given the marked clinical deficits associated with a clinical dementia. COG features only actually reached excellent performance as well (97.6% positive predictive value, 98.0% specificity, 96.4% sensitivity), while using VBM features only yielded markedly lower performances (86.4% positive predictive value, 89.3% specificity, 79.6% sensitivity). Note that the performance metrics in ADNI1 were estimated through cross-validation, and represent an average performance for several models trained on different subsets of ADNI1. We then trained a model on all of ADNI1, and estimated its performance on an independent dataset, ADNI2. Using VCOG predictors, the ADNI1 model reached 92.0% positive predictive value (96.3% specificity, 92.0% sensitivity), when applied on ADNI2, AD vs CN data. Again the performance was comparable with COG predictors only (92.2% positive predictive value, 96.3% specificity, 94.3% sensitivity), and lower performance with VBM features only (57.3% positive predictive value, 79.8% specificity, 56.7% sensitivity). Note that positive predictive value is dependent on the proportion of patients and controls for a given sensitivity and specificity. Since the ADNI2 sample has a substantially smaller proportion of AD subjects compared to ADNI1, the resulting positive predictive value is reduced. When we adjusted the baseline rate of AD subjects in ADNI2 to the same rate in ADNI1, the positive predictive values were 95.2%, 95.3%, and 70.2% for the VCOG, COG, and VBM models respectively.

5.5.2 Identification of easy AD cases for prediction

We tested the stability of the SVM prediction with VCOG features, using random subsamples of ADNI1. We calculated hit probabilities, representing the frequency at which a given subject would be correctly identified as AD, across all subsamples that did not include that subject as part of the training of the model. When implementing this procedure on all of ADNI1, we found that 88% of subjects had a 100% hit probability, which we defined as easy cases. A logistic regression was then trained to predict the easy AD cases. The logistic regression was regularized, and selected a subset of features to perform the prediction. We called the combination of these features a highly predictive signature (HPS) of AD. The VCOG HPS model achieved 99.2% positive predictive value (99.5% specificity, 77.6% sensitivity) in classifying easy AD subjects in ADNI1. These performance scores were estimated by cross-validation of the entire two-stage process (subtyping, traing of SVM, estimation of hit probability, identification of HPS). The hyperparameters of the two-stage model were optimized on classifying pMCI vs sMCI in ADNI1. We next trained a single model on all of ADNI1, which we then applied on a independent sample (ADNI2). The ADNI1 AD VCOG HPS model reached 98.6% positive predictive value (99.5% specificity, 79.5% sensitivity) on ADNI2. Similarly to the SVM results, the VCOG HPS model had similar performance to the COG HPS model (ADNI1: 100% positive predictive value, 100% specificity, 87.3% sensitivity; ADNI2: 98.7% positive predictive value, 99.5% specificity, 88.6% sensitivity), and outperformed the VBM HPS model (ADNI1: 92.3% positive predictive value, 96.1% specificity, 54.6% sensitivity; ADNI2: 65.2% positive predictive value, 91.5% specificity, 33.3% sensitivity), see Figure 5.2a. When adjusted to the same baseline rate of AD subjects as ADNI1, the positive predictive values in ADNI2 were 99.2%, 99.3%, and 76.7% for the VCOG, COG, and VBM HPS models respectively.

The HPS models consistently outperformed the base SVM classifiers with respect to specificity ($p < 0.001$) in the classifications of AD vs CN and pMCI vs sMCI in both ADNI1 and ADNI2, regardless of the features that the models contained. The HPS also had greater positive predictive value ($p < 0.05$) adjusted for a typical prevalence of 33.6% pMCI in a given sample of MCI subjects (Mitchell and Shiri-Feshki, 2009). However, these increases in specificity and positive predictive value for the HPS model came at a significant cost of reduced sensitivity compared to the base classifier, across all models in both ADNI1 and ADNI2 ($p < 0.05$) (Figure 5.2). Note that this shift towards lower sensitivity and higher specificity and positive predictive value could be achieved by adjusting the threshold of the SVM analysis (see ROC analysis, Figure 5.7 in supplementary material).

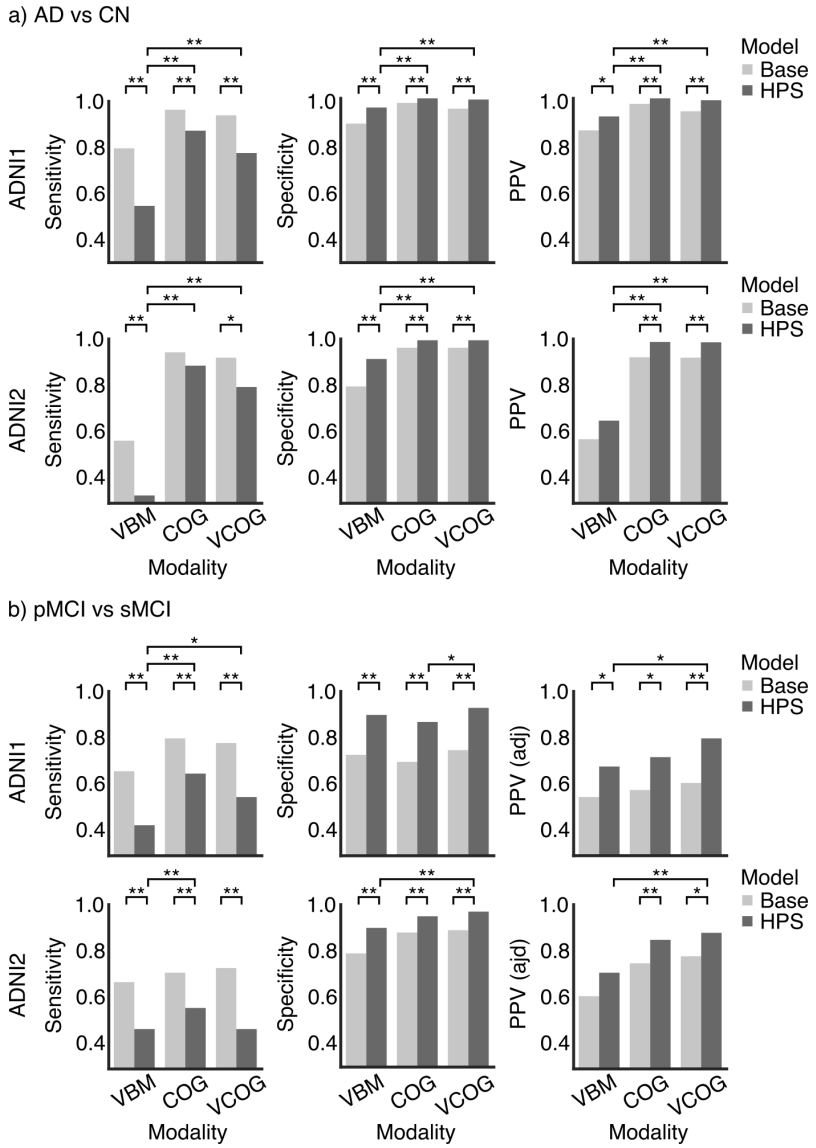

Figure 5.2: Specificity, sensitivity, and positive predictive value (PPV) for the base SVM and highly predictive signature (HPS) classifiers in the classifications of a) patients with AD dementia (AD) and cognitively normal individuals (CN) and b) patients with mild cognitive impairment who progress to AD (pMCI) and stable MCI (sMCI) in ADNI1 and ADNI2. VBM represents the models trained with VBM subtypes, COG represents the models trained with baseline cognitive scores, and VCOG represents the modelsprediction trained with both VBM subtypes and cognition. Significant differences are denoted by $*$ for $p < 0.05$ and $**$ for $p < 0.001$). Positive predictive value was adjusted (PPV (adj)) for a prevalence of 33.6% pMCI in a sample of MCI subjects for both ADNI1 and ADNI2 MCI cohorts.

5.5.3 High confidence prediction of progression to AD dementia

We next assessed whether the HPS AD model trained on ADNI1 could be used to predict pMCI from sMCI (within three years of follow-up). We applied the AD HPS model trained on ADNI1 on 235 subjects within the ADNI1 MCI sample. We stratified the MCI patients into three categories: negative individuals were classified as controls by the baseline SVM model, HPS+ individuals were identified as patients by the AD HPS model, and non-HPS+ individuals were identified as patients by the baseline SVM model, but not by the AD HPS model. When using the full VCOG features, 87 MCI patients were selected as HPS+ in ADNI1, out of which 81 (93.1% positive predictive value) were pMCI within 36 months follow-up. This represented a large, significant increase over the baseline rate of progressors in the entire ADNI1 MCI sample (37.4%) ($p < 0.001$). This was also a significant increase over the SVM's predictions, where 83.9% of subjects that it had labeled as hits were true progressors ($p < 0.001$). When adjusted to a 33.6% baseline rate of progressors, more typical of MCI populations, the positive predictive value of HPS+ for prognosis of dementia was 80.4% (93.2% specificity, 55.1% sensitivity). We replicated these analyses in the MCI sample from ADNI2. Using VCOG features, 32 subjects were identified as HPC+, 26 of which progressed to AD dementia within 36 months follow-up (81.2% positive predictive value, specificity of 96.7%, sensitivity of 47.3%, 87.8% positive predictive value adjusted to a 33.6% baseline rate). This represented a significantly higher prevalence than the 30.6% baseline rate in the entire ADNI2 MCI cohort ($p < 0.001$). This was also a significant increase over the SVM's predictions, where 67.8% of subjects it had labeled as hits were true progressors ($p < 0.001$). As expected, the HPS+ model had lower sensitivity and higher specificity/positive predictive value than the baseline model, both in ADNI1 and ADNI2, and for all three feature sets (VBM, COG and VCOG), see Figure 5.2b. The VCOG features also lead to higher positive predictive value than VBM and COG features taken independently, both in ADNI1 and ADNI2. That increase was large and significant between VCOG and VBM (up to 17%) and marginal, non-significant between VCOG and COG (up to 8%), see Figure 5.2b.

5.5.4 Characteristics of MCI subjects with a highly predictive VCOG signature of AD

HPS+ MCI subjects with the VCOG signature were more likely to be progressors (Figure 5.3a) compared to non-HPS+ subjects and negative subjects (ADNI1: $p < 0.001$; ADNI2: $p < 0.001$). HPS+ MCI subjects were also more likely to be APOE4 carriers (Figure 5.3b) (ADNI1: $p < 0.005$; ADNI2: $p < 0.05$). There was no difference

in sex across the HPS groupings in the MCI subjects of either the ADNI1 or ADNI2 cohorts (Figure 5.3c). This was consistent with the whole sample, where there were equal proportions of progressors across both sexes in each dataset (ADNI1: $\chi^2 = 0.015, p = 0.90$; ADNI2: $\chi^2 = 0.0002, p = 0.99$). The HPS+ class was also significantly enriched for Aβ-positive subjects in ADNI1 ($p < 0.05$). However, this result was not replicated in the ADNI2 MCI subjects (Figure 5.3d). Similarly with tau, we found a significant increase in tau-positive subjects in the HPS+ group of ADNI1 ($p < 0.05$), but not in ADNI2 (Figure 5.3e). We found a significant age difference across the HPS classes in ADNI2 ($F = 5.68, p < 0.005$), where the HPS+ subjects were older than the Negative subjects by a mean of 4.4 years. However, age did not differ across the HPS classes in ADNI1 (Figure 5.3f). Finally, HPS+ subjects had significantly steeper cognitive declines compared to the Non-HPS+ and Negative groups (Figure 5.3g): there were significant interactions between the HPS groupings and time in ADNI1: (HPS+ $\beta = -0.147, t = -7.56, p < 0.001$, Non-HPS+ $\beta = -0.055, t = -2.46, p < 0.05$) and ADNI2 (HPS+ $\beta = -0.194, t = -8.69, p < 0.001$, Non-HPS+ $\beta = -0.072, t = -3.31, p = 0.001$). The HPS+ in ADNI1 and ADNI2 respectively gained 1.8 and 2.3 more points each year on the ADAS13 compared to the Non-HPS+ and Negative groups. Note that higher scores on the ADAS13 represent worse cognitive function.

5.5.5 COG, VBM and VCOG signatures

The COG signature was mainly driven by scores from the ADAS13, which measures overall cognition, ADNI-MEM, which measures memory, and ADNI-EF, which measures executive function (coefficients were 5.49, -4.80 and -2.50 respectively). In this model, sex, age, mean GMV, and TIV contributed very little, relative to the cognitive features (Figure 5.4b). Note that these coefficients should be interpreted as pseudo z-scores as the features had been normalized to zero mean and unit variance.

Almost all grey matter subtypes contributed to the VBM signature. Mean GMV, subtype 1 and subtype 6 had the highest weights in the model (coefficients were -5.07, 4.87, and 3.98 respectively) (Figure 5.4c). Subtype 1 was characterized by reduced relative GMV in the occipital, parietal and posterior temporal lobes. Subtype 6 was characterized by reduced relative GMV in the temporal lobes, notably the medial temporal regions. We had anticipated the larger contribution of these two subtypes as they have been described in previous AD subtyping work (Noh et al., 2014; Hwang et al., 2016; Park et al., 2017; Varol et al., 2017). Diagnosis (CN, sMCI, pMCI, AD) accounted for a substantial amount of variance in these subtype weights (subtype 1: $F = 8.51, p < 0.001$; subtype 2: $F = 34.27, p < 0.001$). Post-hoc t-tests showed AD subjects had significantly higher weights compared to CN (Figure 5.1e), making

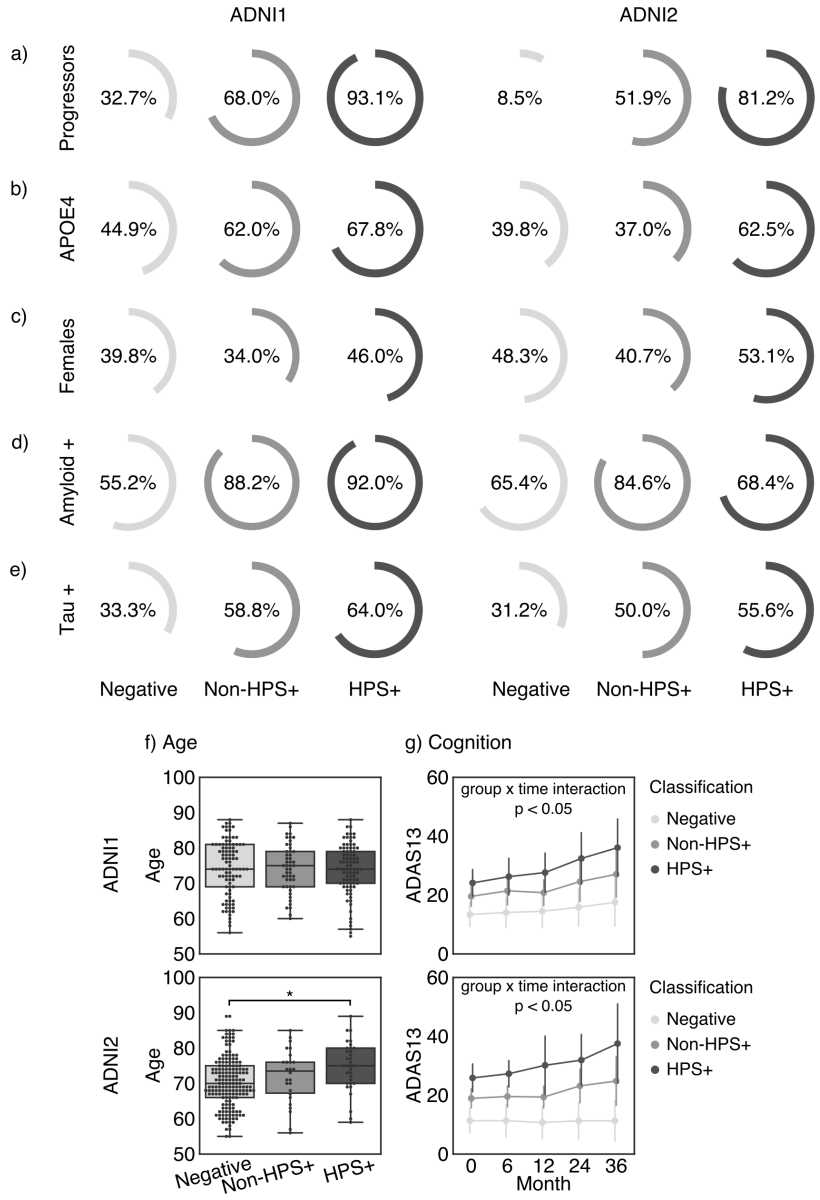

Figure 5.3: Characteristics of MCI subjects with the VCOG signature in ADNI1 and ADNI2. We show the percentage of MCI subjects who a) progressed to dementia, were b) APOE4 carriers, c) female, d) positive for Aβ measured by a cut-off of 192 pg/mL in the CSF (Shaw et al., 2009), and e) positive for tau measured by a cut-off of 93 pg/mL in the CSF (Shaw et al., 2009) in each classification (HPS+, Non-HPS+, and Negative). f) Age and g) cognitive trajectories, measured by the Alzheimer's Disease Assessment Scale - Cognitive subscale with 13 items (ADAS13), across the three classes. Significant differences are denoted by $*$ for family-wise error rate-corrected $p < 0.05$.

these subtypes associated with a diagnosis of AD (subtype 1: $t = 2.88, p < 0.05$; subtype 6: $t = 7.68, p < 0.001$).

The ADAS13, memory (ADNI-MEM) and executive function (ADNI-EF) scores contributed the most to the VCOG HPS signature (coefficients were 6.27, -7.43 and -3.95 respectively, Figure 5.4a). Of the VBM features, subtypes 2, 3 and 7 contributed the most to the signature (coefficients were 1.36, -2.12 and -2.83 respectively). Subtypes 1 and 6, which had the highest positive weights in the VBM model, were given marginal weights in the VCOG model, potentially indicative of redundancy with COG features. Subtype 2, which was associated with AD, was characterized by greater relative GMV in medial parts of the parietal and occipital lobes and the cingulate cortex, but less GMV everywhere else. Subtype 3 was characterized by greater relative GMV in the temporal lobes, insula and striatum. Subtype 7, which was associated with healthy controls, was characterized by greater relative GMV in the parietal, occipital and temporal lobes. Note that the weights for subtypes 3 and 7 were negative in the model, which means that predicted AD and pMCI cases had brain atrophy patterns that were spatially dissimilar to those subtypes.

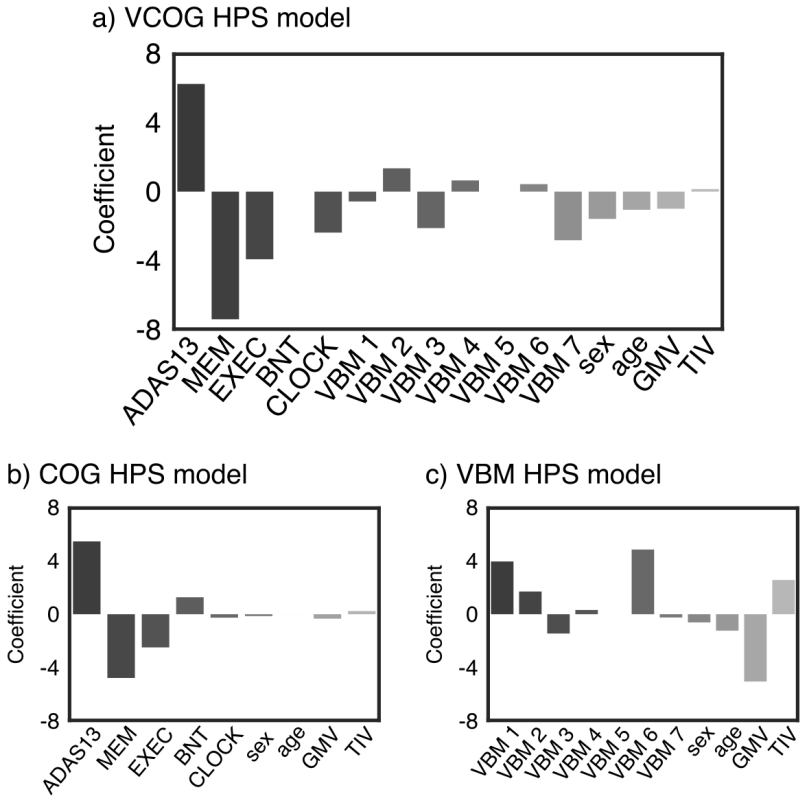

Figure 5.4: Coefficients of the high confidence prediction a) VCOG model, b) COG model, and c) VBM model. ADAS13 = Alzheimer's Disease Assessment Scale - Cognitive, MEM = ADNI-MEM score; EXEC = ADNI-EF score, BNT = Boston Naming Test, CLOCK = clock drawing test, VBM 1-7 = VBM subtype weights, GMV = mean grey matter volume, TIV = total intracranial volume.

5.6 Discussion

We developed a highly precise and specific MRI and cognitive-based model to predict AD dementia. Our two-stage predictive model reached 93.2% specificity and 93.1% positive predictive value (80.4% when adjusted for 33.6% prevalence of progressors) in ADNI1 when classifying progressor vs stable MCI patients (within 3 years follow-up). We replicated these results in ADNI2 where the model reached 96.7% specificity and 81.2% positive predictive value (87.8% adjusted positive predictive value). With respect to specificity and positive predictive value, these results are a substantial improvement over previous work combining structural MRI and cognition on the same prediction task, that have reported up to 76% specificity and 65% positive predictive value (adjusted for 33.6% prevalence of progressors) (Korolev et al., 2016). We also report the highest positive predictive value compared to the current state-of-the-art predictive model using Aβ PET scans, which reported 74% (adjusted) positive predictive value (Mathotaarachchi et al., 2017). Finally, our results are consistent with our previous findings using structural MRI only as well also reproduced our past work which developed a model that optimizes specificity and positive predictive value (Dansereau et al., 2017). Our performance is close to the 90% positive predictive value reported by Dansereau et al. (2017) using a combination of structural and functional MRI measures and a two-stage predictive model, with the limitation of a smaller sample size (N=56 MCI patients) due to limited availability of functional MRI data in ADNI.

An ideal model to predict conversion to AD dementia would have both high sensitivity and high specificity, meaning it would be able to detect all progressor MCI and reject all stable MCI individuals. However, there is a balance between sensitivity and specificity that should be considered depending on the context. In the case of progression to AD dementia, the pathophysiological heterogeneity of clinical diagnosis will prevent highly accurate prediction linking brain features to clinical trajectories. We argue that, faced with heterogeneity, it is necessary to sacrifice sensitivity to focus on a subgroup of individuals with similar brain abnormalities. In a predictive model, one needs to choose a threshold to control the proportions of false positives and false negatives. A receiver operating characteristic (ROC) curve illustrates the trade off between the true positive and false positive rates as the threshold for positivity changes and is often used to determine an optimal threshold. Generally, increasing the threshold leads to higher specificity and lower sensitivity. The high specificity of our two-stage model indeed came at a cost of reduced sensitivity (55.1% in ADNI1 and 47.3% in ADNI2 for classifying pMCI vs sMCI), which is much lower than sensitivity values of 64%-95% reported by other

groups (Dansereau et al., 2017). The two stage procedure, based on prediction stability, did not offer gains compared to a simpler SVM model, if the threshold of the SVM model could be selected *a priori* to match the specificity of the two stage procedure (see ROC curves in Figure 5.7 in supplementary material). The two-stage prediction model offered the advantage of a principled approach to train the prediction model in a high-specificity regime, based on stability. The choice of a level 2 regularized logistic regression also lead to a compact and interpretable subset of features for the HPS. Favoring specificity over sensitivity could be useful in settings where false positives need to be minimized and positive predictive value needs to be high, such as expensive clinical trials. Here, with our HPS VCOG model, we report the high positive predictive values for progression to AD from the MCI stage (up to 87.8%, adjusted for 33.6% prevalence of progressors). Importantly, the proposed HPS+ model used tools that are already widely used by clinicians. The present work could be used as a screening tool for recruitment in clinical trials that target MCI subjects who are likely to progress to dementia within three years. The implementation of an automated selection algorithm could also result in groups of MCI subjects with more homogeneous brain pathology. However, we note that HPS+ subjects did not all present with significant amyloid burden (92.0% and 68.4% of HPS+ subjects in ADNI1 and ADNI2 respectively, Figure 5.3), which means that not all HPS+ individuals are likely to have prodromal AD, even when progressing to AD dementia.

When we trained our model with cognitive features only, tests for general cognition, memory, and executive function were chosen as the strongest predictors of AD dementia. Our COG HPS model thus supports previous research that reported general cognition, memory, and executive function as important neuropsychological predictors of dementia (Dickerson et al., 2007; Belleville et al., 2017; Tabert et al., 2006; Korolev et al., 2016). Compared to the state-of-the-art multi-domain cognition-based predictive model, which reported 87.1% specificity and 81.8% positive predictive value (77.5% when adjusted to 33.6% pMCI prevalence) (Belleville et al., 2014), our COG HPS model achieved similar performance reaching between 87.5%-95% specificity and 72.3%-85.1% (adjusted) positive predictive value. As general cognition was the strongest feature in our model to predict progression, this supports previous findings that MCI patients with deficits across multiple domains are at the highest risk for dementia (Tabert et al., 2006; Mitchell et al., 2009).

For our VBM model, we extracted a number of gray matter atrophy subtypes that recapitulated previously reported subtypes, namely the medial temporal lobe and parietal dominant subtypes (Noh et al., 2014; Hwang et al., 2016; Park et al., 2017; Varol et al., 2017), which were associated strongly with a diagnosis of AD

dementia. Weights for the parietal dominant and medial temporal lobe subtypes (Subtypes 1 and 6 from Figure5.1b, respectively) contributed substantially to the highly predictive signature in the VBM model. The atrophy pattern of subtype 6 is spatially similar to the spread of neurofibrillary tangles in Braak stages III and IV (Braak and Braak, 1991), which may support previous findings that tau aggregation mediates neurodegeneration (Spillantini and Goedert, 2013). The contributions of the parietal dominant and medial temporal lobe subtypes in the VBM model are also in line with previous works, which have reported that cortical thickness and volumes of the medial temporal lobes, inferior parietal cortex, and precuneus are strong predictors of progression to dementia (Korolev et al., 2016; Eskildsen et al., 2013).

When combined with cognitive tests in the VCOG model, the structural subtypes were given marginal weights. This suggests some redundancy between atrophy and cognition, and that cognitive features have higher predictive power than structural features in the ADNI MCI sample. This conclusion is consistent with the observation that the COG model significantly outperformed the VBM model, similar to previous work (Korolev et al., 2016). Although cognitive markers were stronger features, the VCOG model assigned large negative weights for the structural subtypes 3, which showed greater relative GMV in the temporal lobes, and 7, which showed greater relative GMV in the parietal, occipital, and temporal lobes. This means that these features were predictive of stable MCI in the VCOG model, in line with previous work showing that atrophy in these regions is predictive of progression to dementia (Korolev et al., 2016; Eskildsen et al., 2013). Furthermore, we demonstrated that combining MRI data with cognitive markers significantly improves upon a model based on MRI features alone. This result is again in line with the literature (Korolev et al., 2016; Moradi et al., 2015), yet was shown for the first time for a model specifically trained for high positive predictive value. Note that in the current study, the predictive model was trained exclusively on images acquired on 1.5T scanners from ADNI1. Good generalization to ADNI2 with 3T scanners demonstrate robustness of imaging structural subtypes across scanner makes.

While the VCOG model identified the majority of MCI subjects who were highly likely to progress to dementia, there were sizable groups of subjects who were uniquely identified as progressors by the unimodal signatures. For example, out of all the MCI subjects that were picked up by the three HPS models, 12.9% and 35.3% of subjects from ADNI1 and ADNI2 respectively were identified as progressors exclusively by the VBM HPS model (Supplementary Figure 5.9). Although these subjects presented an AD-like brain atrophy pattern, they may exert compensatory mechanisms or have a high cognitive reserve that may allow them to maintain a higher level

of cognitive function than subjects who were also selected by the COG signature. Out of all the HPS+ MCI subjects across the three models, the COG HPS model also exclusively labeled 20.4% and 10.8% of subjects from ADNI1 and ADNI2 respectively as progressors (Supplementary Figure 5.9). As these subjects were not labeled by the VBM model, they may represent subjects who have patterns of neurodegeneration that are atypical of AD or subjects who have a non-AD pathology. Structural MRI and cognition therefore provide complementary information about AD risk, and it appears that brain atrophy can add to the predictive value of cognitive markers. Our results converge with clinical practices that utilize both neuropsychological testing and brain imaging to inform diagnosis and prognosis, although our approach has the added benefits of being data-driven, automated and optimized for high positive predictive value.

The VCOG highly predictive signature might reflect a late disease stage. We looked at the ratio of early to late MCI subjects in the ADNI2 sample (note that ADNI1 did not have early MCI subjects). Of the MCI subjects identified as HPS+ by the VCOG model, 84.4% were late MCI subjects, compared to a rate of 34.9% of late MCI subjects in the entire ADNI2 MCI sample (Supplementary Figure 5.6. In both ADNI1 and ADNI2, the VCOG HPS+ subjects had on average the worst baseline cognition and the worst cognitive decline over time (see Figure 5.3g and Supplementary Figure 5.5). The HPS+ subjects in ADNI1 also had significantly greater amounts of Aβ and tau pathology compared to the whole sample, although these findings were not replicated in ADNI2 (Figure 5.3d,e). Given that the majority of HPS+ subjects were late MCI subjects and had worse cognition, the two-stage predictive model appears to be effective at identifying individuals at later disease stages who are at the cusp of progressing to dementia. This result is to be expected since the model was built on cognitive variables that are designed to separate cognitively normal and individuals with dementia at certain cut-offs. This approach may not be optimal for early detection of future cognitive decline. Training a model to classify MCI progressors and non-progressors to dementia could be done in order to capture future progressors in earlier preclinical stages (e.g. early MCI). Finally, we focused on structural MRI and neuropsychological batteries as features in our models due to their wide availability and established status as clinical tools. However, we believe adding other modalities such as PET imaging, CSF markers, functional MRI, genetic factors, or lifestyle factors could result in higher predictive power, especially at earlier preclinical stages of AD.

5.7 Conclusion

In summary, we developed a highly precise and specific model to predict progression from MCI to AD dementia within a time span of three years. Using structural MRI and neuropsychological markers, we trained a model to recognize an anatomical and cognitive signature that is highly predictive of AD dementia by learning the difference between AD patients and controls. We validated the model in three independent samples, a second sample of AD patients and controls, and two separate cohorts that contained both stable MCI patients and MCI patients who progressed to dementia. The model was able to predict progression to dementia in MCI patients with up to 93.1% positive predictive value and up to 96.7% specificity. Our model could potentially improve subject selection in clinical trials and identify individuals at a higher risk of AD dementia for early intervention in clinical settings.

5.8 Acknowledgements

We would like to thank Sylvia Villeneuve, Alexa Pichet-Binette and Jacob Vogel for providing us with data to help start our preliminary analyses for this project. We would like to thank Hien Nguyen for advising us on parts of the statistical analyses.

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 DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; 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.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health[1]. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego.

[1] www.fnih.org

ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The computational resources used to perform the data analysis were provided by Compute Canada[2]. This project was funded by NSERC grant number RN000028 and the Canadian Consortium on Neurodegeneration in Aging (CCNA)[3], through a grant from the Canadian Institutes of Health Research and funding from several partners including SANOFI-ADVENTIS R&D. AT is supported by a bursary from the Centre de recherche de l'institut universitaire de gériatrie de Montréal and the Courtois foundation. CD is supported by a salary award from the Lemaire foundation and Courtois foundation. PB is supported by a salary award from "Fonds de recherche du Québec – Santé" and the Courtois foundation.

5.9 Supplementary material

5.9.1 Supplementary results

Comparison of COG, VBM and VCOG HPS+ subjects

We found substantial overlap of subjects labeled as HPS+ in the MCI cohorts across the VBM, COG and VCOG signatures (Supplementary Figure 5.9). There were very few subjects that were labeled as HPS+ exclusively by the VCOG signature. As to be expected, the majority of subjects labeled as HPS+ by the VCOG signature (ADNI1: 97.7%; ADNI2: 100%) were also labeled as HPS+ by either the VBM only or COG only signatures or both. Of the subjects that were labeled as HPS+ by the VBM only signature, 23.6% and 55.2% in ADNI1 and ADNI2 respectively were identified exclusively by the VBM HPS. There were relatively few subjects (7 and 2 subjects in ADNI1 and ADNI2 respectively) that were captured by VBM and VCOG but missed by the COG HPS. The COG HPS actually identified the majority of all HPS+ subjects across the three signatures (ADNI1: 106 of 132 total subjects, ADNI2: 40 of 65 total subjects). From Supplementary Figure 5.9, we can see that the VCOG HPS acts as a refinement of the COG signature, as the VCOG HPS captures a subset of subjects that were labeled by the COG HPS.

Out of the HPS+ subjects labeled by all three signatures, 97.9% and 93.7% from ADNI1 and ADNI2 respectively progressed to dementia (Supplementary Table 5.2). These subjects had worse cognition based on the MMSE and higher proportions of APOE4 carriers, Aβ positive and tau positive individuals, compared to the baseline rates in all MCI subjects. Of the HPS+ subjects who were labeled only by the VBM

[2] https://computecanada.org/
[3] http://ccna-ccnv.ca

model, 70.6% and 43.4% from ADNI1 and ADNI2 respectively were progressors. This group of subjects had less Aβ and tau positive individuals compared to the baseline rates. Of the HPS+ subjects who were labeled only the COG model, 70.4% and 57.1% from ADNI1 and ADNI2 respectively progressed to dementia. This group appeared to have a greater proportion of Aβ positive individuals compared to the baseline rates in both ADNI1 and ADNI2 cohorts. The majority of these COG HPS+ subjects were also male. Given the distinct characteristics among the exclusively COG, exclusively VBM, and VCOG HPS+ subjects, these groups may represent subgroups with different risks for AD dementia. As it appears that a greater proportion of pMCI is captured when cognitive and structural MRI features are combined, these findings may support joining multiple modalities together in order to achieve higher positive predictive value. However, these results are qualitative and of an exploratory nature due to low sample sizes.

5.9.2 Supplementary figures

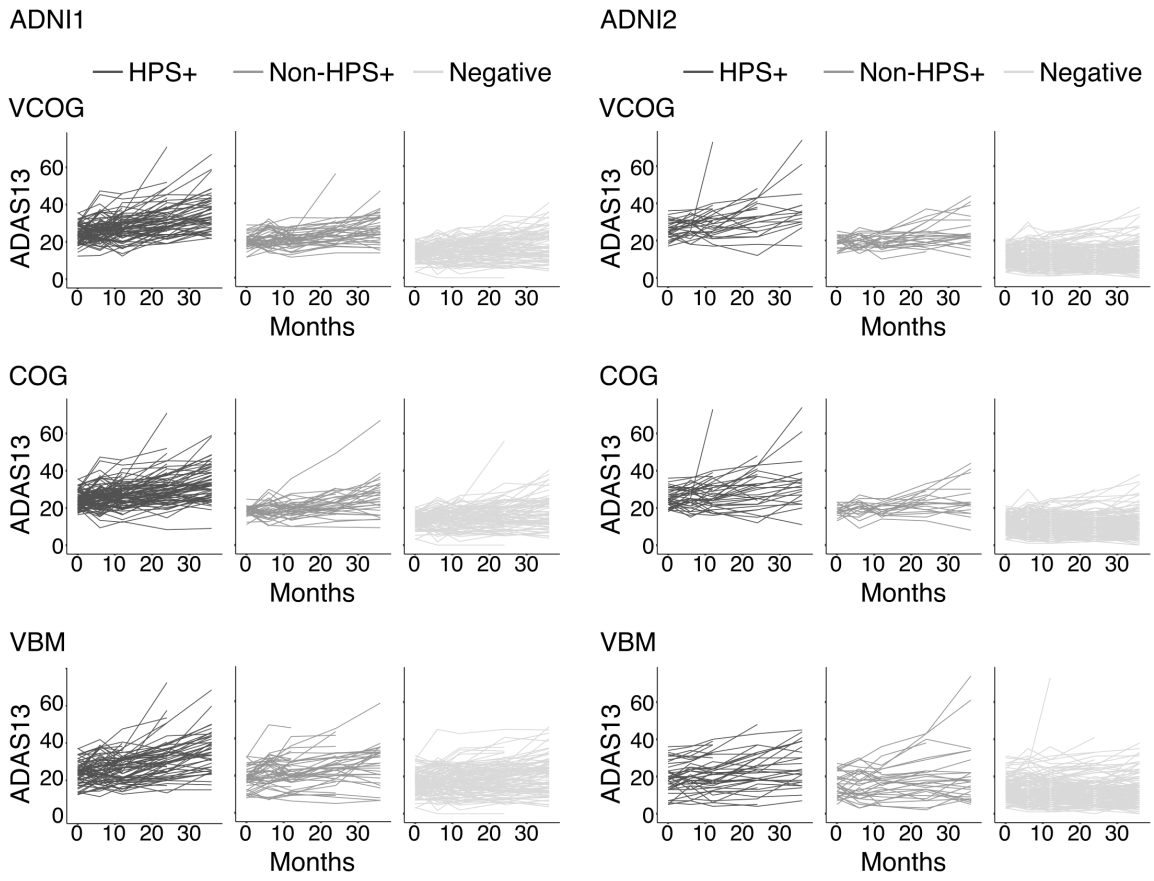

Figure 5.5: Cognitive trajectories for individual MCI subjects in ADNI1 and ADNI2 grouped by HPS classifications (HPS+, Non-HPS+, Negative) by the VCOG, COG, and VBM high confidence prediction models.

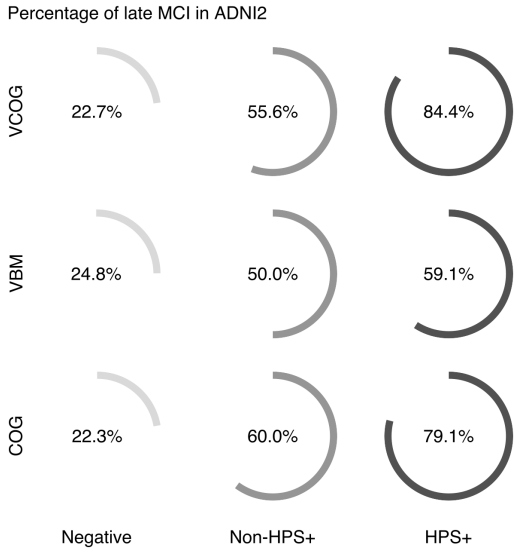

Figure 5.6: Percentage of late MCI ADNI2 subjects within each HPS grouping (HPS+, Non-HPS+, Negative) across each highly predictive model (VCOG, VBM, COG). In each model, there was a significantly greater proportion of late MCI subjects in the HPS+ class compared to the Negative class (VCOG: $\chi^2 = 51.0, p < 0.001$; VBM: $\chi^2 = 21.8, p < 0.001$; COG: $\chi^2 = 59.9, p < 0.001$).

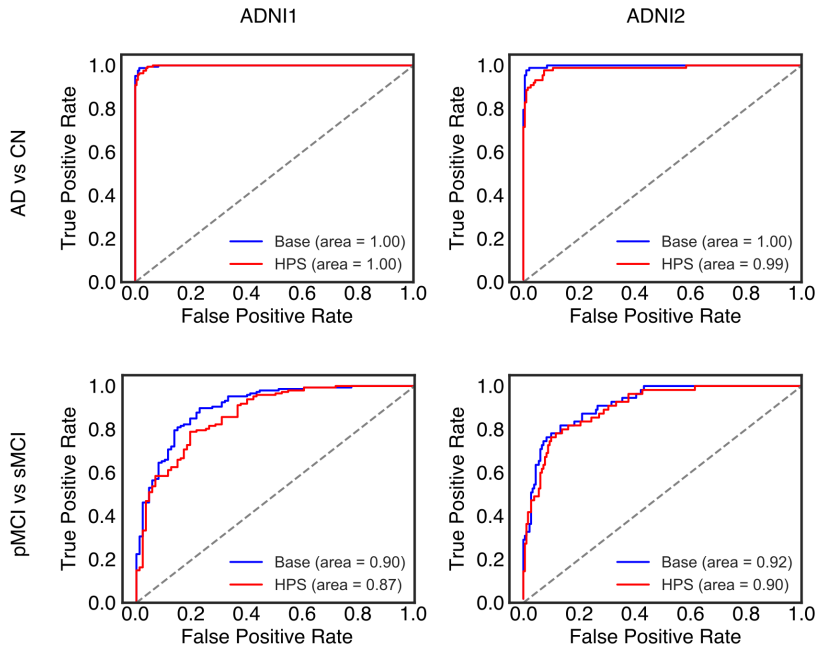

Figure 5.7: Receiver operating curves (ROC) for each classification (AD vs CN; pMCI vs sMCI) and model (Base: traditional SVM; HPS: two-stage highly predictive signature) in both ADNI1 and ADNI2 cohorts. The ROC plots are shown for models using the VCOG predictors.

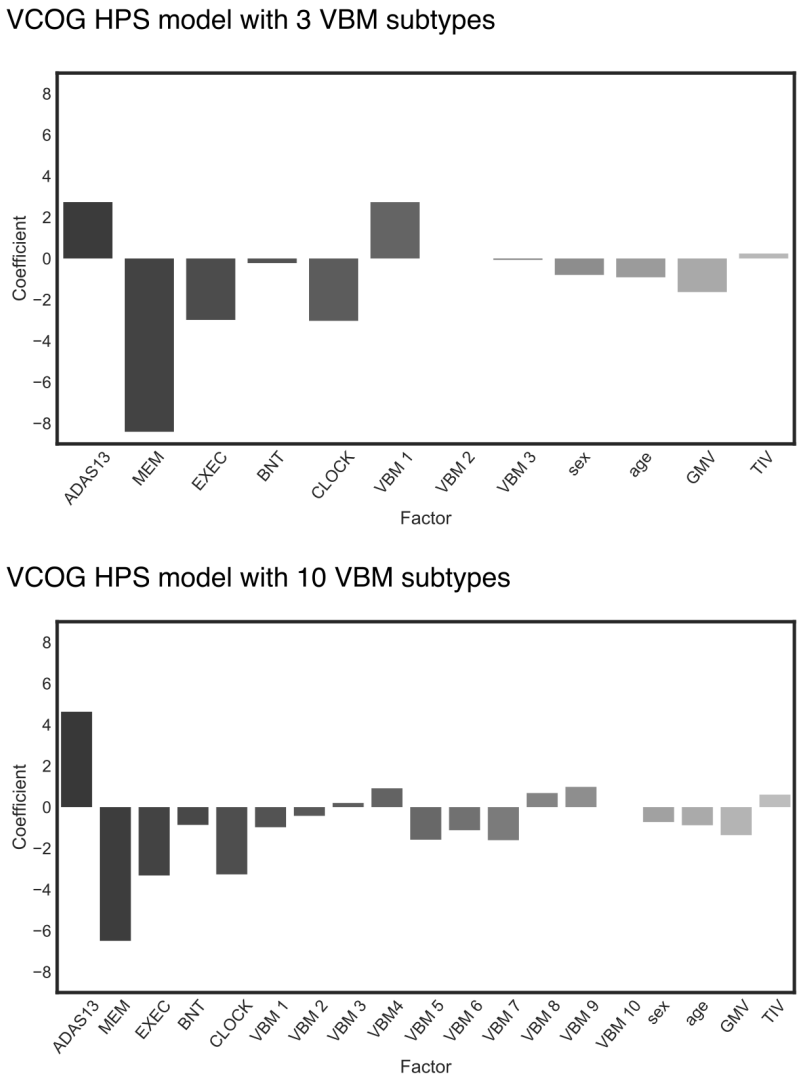

Figure 5.8: Coefficients of factors in the VCOG HPS models for a model featuring 3 VBM subtypes and a model featuring 10 VBM subtypes.

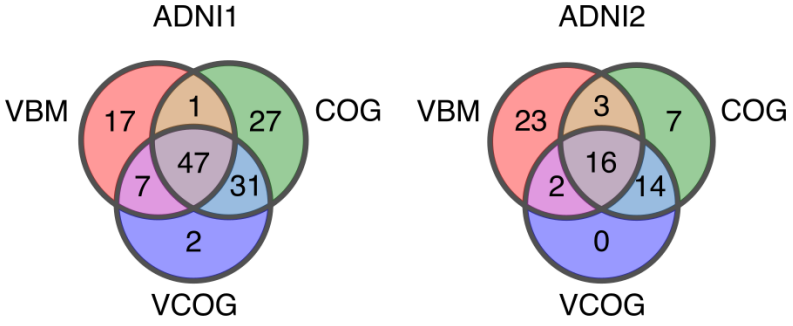

Figure 5.9: Venn diagram depicting the number of MCI subjects labeled as HPS+ by the VBM, COG, and VCOG HPS models in ADNI1 and ADNI2.

5.9.3 Supplementary tables

Table 5.2: Characteristics of HPS+ subjects from the VBM, COG, and VCOG signatures in ADNI1 and ADNI2 MCI cohorts

ADNI1	**Baseline**	**VBM** ~~COG~~ ~~VCOG~~	~~VBM~~ **COG** ~~VCOG~~	~~VBM~~ ~~COG~~ **VCOG**	**VBM** **COG** ~~VCOG~~	**VBM** ~~COG~~ **VCOG**	~~VBM~~ **COG** **VCOG**	**VBM** **COG** **VCOG**
pMCI %	62.6	70.6	70.4	100	100	57.1	93.5	97.9
APOE4 %	57.0	58.8	70.4	100	100	71.4	71.0	63.8
Aβ+ %	76.1	50.0	91.7	n/a	100	100	88.9	91.7
tau+ %	50.7	0.0	83.3	n/a	0.0	50.0	66.7	66.7
Female %	40.8	41.2	29.6	50.0	0.0	42.9	61.3	36.2
Age	74.1±7.2	74.4±5.5	73.9±6.0	80.0±6.2	79.3	69.3±4.1	73.9±7.0	74.4±7.3
Education	15.7±2.9	17.1±2.1	15.7±3.3	15.0±1.4	20.0	16.0±2.6	14.3±3.1	15.9±2.9
MMSE	27.0±1.8	27.8±1.4	27.6±1.2	26.5±0.7	29.0	25.4±0.5	25.9±1.5	26.0±1.6
ADNI2	**Baseline**	**VBM** ~~COG~~ ~~VCOG~~	~~VBM~~ **COG** ~~VCOG~~	~~VBM~~ ~~COG~~ **VCOG**	**VBM** **COG** ~~VCOG~~	**VBM** ~~COG~~ **VCOG**	~~VBM~~ **COG** **VCOG**	**VBM** **COG** **VCOG**
pMCI %	30.6	43.5	57.1	n/a	33.3	0.0	78.6	93.7
APOE4 %	42.6	47.8	42.9	n/a	33.3	50.0	57.1	68.7
Aβ+ %	69.0	75.0	100	n/a	100	100	55.6	77.8
tau+ %	39.7	42.9	25.0	n/a	50.0	0.0	62.5	55.6
Female %	48.1	39.1	14.3	n/a	33.3	50.0	57.1	50.0
Age	70.7±7.3	71.4±5.9	74.8±6.6	n/a	67.0±13.7	73.6±2.5	71.7±8.6	77.1±4.8
Education	16.4±2.6	16.3±2.6	16.9±2.5	n/a	15.0±3.6	19.0±1.4	15.6±2.3	15.9±2.8
MMSE	28.2±1.7	28.3±1.5	28.0±1.3	n/a	28.7±1.5	27.5±2.1	26.7±1.8	26.1±1.7

The Baseline column represents values for all MCI subjects.
Age and education are presented in years (mean ± standard deviation).
Aβ and tau CSF measures were available for approximately one third of all subjects across both cohorts.

Table 5.3: pMCI vs sMCI performance metrics for VCOG HPS models with different number of VBM subtypes

ADNI1	**3 clusters**	**7 clusters**	**10 clusters**
Specificity	0.8636	0.9310	0.8864
Sensitivity	0.5578	0.5510	0.5510
Positive predictive value (adjusted)	0.6743	0.8035	0.7105
Accuracy	0.6723	0.6936	0.6766
ADNI2	**3 clusters**	**7 clusters**	**10 clusters**
Specificity	0.9556	0.9667	0.9556
Sensitivity	0.4727	0.4727	0.4364
Positive predictive value (adjusted)	0.8433	0.8777	0.8324
Accuracy	0.8425	0.8511	0.8340

Positive predictive value was adjusted for a 33.6% prevalence of pMCI subjects.

Chapter 6

Discussion

The studies in this thesis explored heterogeneity in structural and functional brain organization with MRI in order to advance neuroimaging as a tool to measure biomarkers for Alzheimer's disease. Specific contributions and future works derived from chapters 3, 4, and 5 will be discussed below.

6.1 Contributions

6.1.1 Functional connectivity as a robust biomarker

In chapter 3, we reported functional connections within the default mode network and cortico-striatal-thalamic loop that were consistently diminished in MCI patients compared to controls in several independent datasets. This result suggests that these functional connections could be robust biomarkers of prodromal AD in spite of variability from data acquisition. It is worth noting though that when we examined each sample independently, there were few significant findings and that the majority of effects found within each sample did not appear to be reproducible due to the substantial heterogeneity among the samples. As a consequence, after pooling the datasets together, we found small to medium effect sizes and our sample size estimates recommend at least 140 to over 600 subjects to achieve adequate statistical power. Overall, the major contribution from this study was identifying functional connections which may be candidate biomarkers for AD. Functional parcellations derived from this study are described in chapter 4, and we have also publicly released the statistical maps so that they can be directly used in future studies.

Over the course of my thesis, I contributed to other studies that also examined the relationship between resting-state functional connectivity and the AD spectrum. This included a meta-analysis on the functional connections that were reported to

be affected in AD and MCI patients in the literature (see Badhwar et al. (2017) in Appendix A). Although there was consensus across the literature about decreased connectivity within the default mode network in patient groups, we also noted high variability between reported regions. Given the heterogeneity seen in chapter 3 and in Badhwar et al. (2017), I also contributed to a study that investigated subtypes of functional connectivity patterns and whether certain subtypes were associated with a higher risk of AD (see Orban et al. (2017) in Appendix B). In this latter study, we found subtypes of connectivity in default mode, salience, and limbic networks were associated with cognitive impairments, and limbic subtypes were correlated with amyloid pathology.

6.1.2 Automated diagnosis and precision medicine

In chapter 5, we combined machine learning with MRI and neuropsychological evaluations to automatically predict future progression to AD dementia in individuals with MCI. We improve upon current diagnostic standards in terms of positive predictive value and specificity by identifying a homogeneous subset of individuals who are at a particularly high risk for AD dementia. In doing so, we found specific patterns, or subtypes, of brain atrophy that were associated with cognitive impairment related to AD. This study also assessed the predictive value of cognitive assessments and structural MRI and evaluated the complementarity of these two modalities for diagnostic and prognostic purposes.

I also contributed to Dansereau et al. (2017) (see Appendix C), where this study introduced the two-stage predictive algorithm that was used in chapter 5. This algorithm optimizes for high positive predictive value and specificity by learning cases that are easy to classify. In Dansereau et al. (2017), we also found that the combination of functional connectivity subtypes and atrophy subtypes yielded better performance to predict future progression to AD dementia than either modality on its own.

Our results, especially those of chapter 5, recapitulate the work flow of clinicians to identify certain patterns of brain atrophy in order to supplement the diagnostic information offered by neurospychological evaluations. The advantage of our work is that this diagnostic process can be automated. Furthermore, by characterizing the heterogeneity within populations along the AD spectrum, we were able to identify, with high specificity, subgroups with particular brain characteristics that were highly predictive of incipient AD dementia. Examining subtypes of brain organization represents an incremental step towards customized healthcare that is tailored to an individual. Lastly, optimizing for high specificity may be beneficial for future diagnostic algorithms for AD given that the majority of individuals, even those with

preclinical signs (e.g. amyloid burden, neurodegeneration), will never develop AD dementia in their lifetime (Brookmeyer and Abdalla, 2018). Overall, the studies from chapter 5 and Dansereau et al. (2017) make advances toward automated diagnosis and precision medicine.

6.1.3 Subtypes of cortical atrophy

As previously mentioned, we reported the presence of numerous distinct patterns of cortical atrophy in the AD spectrum in chapter 5. We described four patterns that were associated with a diagnosis of AD dementia: 1) predominantly temporal lobe atrophy, 2) grey matter loss in the parietal and occipital lobes, 3) reduced grey matter in default mode regions, and 4) diffuse atrophy. This work reveals the heterogeneous nature of neurodegeneration and corroborates previous works that have explored structural subtypes in AD (Noh et al., 2014; Hwang et al., 2016; Park and Friston, 2013; Varol et al., 2017; Zhang et al., 2016). More specifically, we were able to reproduce the "medial temporal dominant", "parietal dominant", and "diffuse" atrophy subtypes that have been described in the literature, which can be seen in Figure 2.3 from Hwang et al. (2016).

Together, these works challenge the notion that AD-related neurodegeneration occurs in a stereotyped manner that always starts in the temporal lobe and spreads throughout the rest of the cortex. Although, it is notable that individuals expressing structural subtypes with temporal lobe atrophy are more likely to progress to AD dementia, as we demonstrated in both chapter 5 and Dansereau et al. (2017) (see Appendix C). Furthermore, the presence of distinct subtypes of atrophy may shed some light on the heterogeneity in functional connectivity differences between MCI patients and controls across independent cohorts that we saw in chapter 3. It would be very interesting to explore how much the variance in brain structure may impact the variance in functional connectivity.

One reason behind this apparent heterogeneity in cortical atrophy could be the spread of tau, given that the hyperphosphorylation of tau plays a key role in many neurodegenerative diseases (Iqbal, Liu, and Gong, 2016). Subtypes of brain atrophy actually correspond quite well to neuropathological subtypes of AD based on tau deposition from Murray et al. (2011). It is thus possible that these subtypes of atrophy are driven by the abnormal hyperphosporylation and spreading of tau from different epicenters in the brain. Future works should consider examining these structural subtypes along with tau imaging to decipher a potential mechanism for the variance in neurodegeneration.

6.1.4 Data sharing and reproducibility

This thesis also makes minor contributions to open science. The majority of the data used in these studies was publicly available. In cases where there was a scarcity of data, as was the case for the studies with resting-state fMRI in chapters 3 and 4, we pooled data collected by collaborators. Given the small effect sizes we discovered with respect to functional connectivity differences between MCI patients and controls, we must note that the analyses in chapter 3 would not have been possible without the amalgamation of several datasets in order to have a sufficient sample size. This underscores the importance of data sharing in the scientific community in order to obtain reliable results. Through chapter 4, we publicly released the functional parcellations and statistical maps we derived from our analyses in chapter 3. Not only will this sharing of data derivatives help other researchers reproduce our results, but our functional clusters could be used as target regions in future studies. All of the studies presented in this thesis were conducted with free open-source software and we have publicly shared the code so that other researchers could use it and reproduce our findings.

6.2 Future works

6.2.1 Multimodal models

As our results demonstrate that combining multiple modalities (e.g. cognition and MRI) in a predictive model can be more powerful than models with single modalities (e.g. MRI only, cognition only), a natural extension of our work would be to add more modalities to our models. The addition of variables such as molecular biomarkers of AD (i.e. Aβ and tau CSF or PET images), genetics, or lifestyle factors may enable the development of a more comprehensive model for prediction. As MRI is a non-invasive and relatively inexpensive imaging tool (compared to PET) that is already widely used in clinical settings, I am personally interested in developing models incorporating different MRI modalities (e.g. fMRI, anatomical T1-weighted images, white matter tractography) to better understand the neural dysfunction that occurs in AD and to create a biomarker that could be deployed faster in clinical contexts.

6.2.2 Applications to clinical trials

Chapters 3 and 4 of this thesis generated endpoints that could be used in clinical trials for AD. Functional connectivity differences between MCI patients and controls

in default mode and cortico-strial-thalamic regions were robust to site variability, which may make for biomarkers that are well-suited for a clinical trial.

The predictive model from chapter 5 can be directly applied in a clinical trial as a tool for subject selection. Current methods for enrichment of individuals who are at high risk of AD dementia include amyloid measurement (via PET or CSF), APOE4 carriage, and family history of AD. While the presence of amyloid, the APOE4 allele, and a family history of AD are all important risk factors to progression to AD dementia, they are not particularly specific given that a substantial proportion of individuals who are positive for any of these risk factors will never experience AD dementia in their lifetimes. An advantageous feature of the predictive model from chapter 5 is that it is highly specific (in this case, it is specific for selecting patients with MCI who will progress to AD dementia). For example, if a clinical trial is evaluating a treatment in prodromal individuals with the aim to prevent dementia, the trial needs to have criteria that will correctly select subjects who will indeed progress to dementia. This is where having a model with high specificity (and thus high positive predictive value) can be beneficial. Optimizing for high specificity for subject selection may have a large impact on whether a drug will be deemed effective. If the selection criteria are not specific enough, many subjects who do not have the targeted disease may erroneously enter the trial and obscure the intended effects of the drug. Another advantage of our model is that we used features that were easily accessible (i.e. neuropsychological assessements and MRI). If our model could be used as a screening tool in lieu of amyloid PET, the financial savings would be substantial.

6.3 Conclusion

In summary, this thesis explored heterogeneity in patterns of neurodegeneration and functional connectivity in order to develop signatures with MRI that could be predictive of individuals' cognitive trajectories, namely future progression to AD dementia. The results presented here demonstrate how different sources of heterogeneity, ranging from data acquisition to individual variability, may impede the development of a precise or reliable biomarker. By exploiting this heterogeneity, we tested the robustness of functional connectivity as an early biomarker of AD and we developed a signature with high specificity and positive predictive value at predicting incipient AD dementia. The results from this thesis may be used for future biomarker discovery and clinical trial applications.

Bibliography

Ad-Dab'bagh, Y et al. (2006). "The CIVET image-processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research". In: *Proceedings of the 12th annual meeting of the organization for human brain mapping*, p. 2266.

Aggarwal, N T et al. (2005). "Mild cognitive impairment in different functional domains and incident Alzheimer's disease." In: *Journal of neurology, neurosurgery, and psychiatry* 76.11, pp. 1479–1484.

Aggarwal, Neelum T et al. (2006). "Motor dysfunction in mild cognitive impairment and the risk of incident Alzheimer disease." In: *Archives of Neurology* 63.12, pp. 1763–1769.

Agosta, Federica et al. (2012). "Resting state fMRI in Alzheimer's disease: beyond the default mode network". In: *Neurobiology of Aging* 33.8, pp. 1564–1578.

Aisen, Paul S, Bruno Vellas, and H Hampel (2013). "Moving towards early clinical trials for amyloid-targeted therapy in Alzheimer's disease." In: *Nature Reviews Drug Discovery* 12.4, pp. 324–324.

Albert, Marilyn S et al. (2011). "The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." In: *Alzheimer's & dementia : the journal of the Alzheimer's Association* 7.3, pp. 270–279.

Alzheimer's Association (2018). "2018 Alzheimer's disease facts and figures". In: *Alzheimer's & Dementia* 14.3, pp. 367–429.

Ashburner, John (2007). "A fast diffeomorphic image registration algorithm". In: *NeuroImage* 38.1, pp. 95–113.

Ashburner, John and Karl J Friston (2000). "Voxel-Based Morphometry—The Methods". In: *NeuroImage* 11.6, pp. 805–821.

Badhwar, AmanPreet et al. (2017). "Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis." In: *Alzheimers Dementia Diagnosis, Assessment Disease Monitoring* 8, pp. 73–85.

Bai, Feng et al. (2009). "Abnormal resting-state functional connectivity of posterior cingulate cortex in amnestic type mild cognitive impairment". In: *Brain Research* 1302.C, pp. 167–174.

Bateman, Randall J et al. (2012). "Clinical and Biomarker Changes in Dominantly Inherited Alzheimer's Disease". In: *The New England journal of medicine* 367.9, pp. 795–804.

Beach, Thomas G et al. (2012). "Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010." In: *Journal of Neuropathology & Experimental Neurology* 71.4, pp. 266–273.

Bekris, Lynn M et al. (2010). "Genetics of Alzheimer disease." In: *Journal of Geriatric Psychiatry and Neurology* 23.4, pp. 213–227.

Bellec, P et al. (2011). "A neuroimaging analysis kit for Matlab and Octave". In: *Proceedings of the 17th International Conference on Functional Mapping of the Human Brain*.

Bellec, Pierre (2013). "Mining the Hierarchy of Resting-State Brain Networks: Selection of Representative Clusters in a Multiscale Structure". In: *Pattern Recognition in Neuroimaging (PRNI), 2013 International Workshop on*, pp. 54–57.

Bellec, Pierre et al. (2006). "Identification of large-scale networks in the brain using fMRI." In: *NeuroImage* 29.4, pp. 1231–1243.

Bellec, Pierre et al. (2010). "Multi-level bootstrap analysis of stable clusters in resting-state fMRI". In: *NeuroImage* 51.3, pp. 1126–1139.

Bellec, Pierre et al. (2012). "The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows." In: *Frontiers in neuroinformatics* 6, p. 7.

Bellec, Pierre et al. (2015). "Impact of the resolution of brain parcels on connectome-wide association studies in fMRI." In: *NeuroImage* 123, pp. 212–228.

Belleville, Sylvie et al. (2014). "Predicting decline in mild cognitive impairment: A prospective cognitive study." In: *Neuropsychology* 28.4, pp. 643–652.

Belleville, Sylvie et al. (2017). "Neuropsychological Measures that Predict Progression from Mild Cognitive Impairment to Alzheimer's type dementia in Older Adults: a Systematic Review and Meta-Analysis." In: *Neuropsychology Review* 27.4, pp. 328–353.

Benjamini, Yoav and Yosef Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1, pp. 289–300.

Bennett, D A et al. (2006). "Neuropathology of older persons without cognitive impairment from two community-based studies." In: *Neurology* 66.12, pp. 1837–1844.

Biomarkers Definitions Working Group et al. (2001). "Biomarkers and surrogate endpoints: preferred definitions and conceptual framework". In: *Clinical pharmacology & therapeutics* 69.3, pp. 89–95.

Blennow, Kaj et al. (2010). "Cerebrospinal fluid and plasma biomarkers in Alzheimer disease". In: *Nature Reviews Neuroscience* 6.3, pp. 131–144.

Bossers, Koen et al. (2010). "Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease." In: *Brain* 133.Pt 12, pp. 3699–3723.

Braak, Heiko and Eva Braak (1990). "Alzheimer's Disease: Striatal Amyloid Deposits and Neurofibrillary Changes". In: *Journal of Neuropathology & Experimental Neurology* 49.3, p. 215.

Braak, Heiko and Eva Braak (1991). "Neuropathological stageing of Alzheimer-related changes." In: *Acta neuropathologica* 82.4, pp. 239–259.

Brookmeyer, Ron and Nada Abdalla (2018). "Estimation of lifetime risks of Alzheimer's disease dementia using biomarkers for preclinical disease". In: *Alzheimer's & Dementia*, pp. 1–8.

Buckner, R L et al. (2005). "Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory." In: *Journal of Neuroscience* 25.34, pp. 7709–7717.

Busse, A et al. (2006). "Mild cognitive impairment: long-term course of four clinical subtypes." In: *Neurology* 67.12, pp. 2176–2185.

Carmichael, Owen et al. (2013). "Coevolution of brain structures in amnestic mild cognitive impairment." In: *NeuroImage* 66, pp. 449–456.

Chao, Linda L et al. (2009). "Patterns of cerebral hypoperfusion in amnestic and dysexecutive MCI." In: *Alzheimer Disease & Associated Disorders* 23.3, pp. 245–252.

Chertkow, Howard et al. (2010). "Multilingualism (But Not Always Bilingualism) Delays the Onset of Alzheimer Disease: Evidence From a Bilingual Community". In: *Alzheimer Disease & Associated Disorders* 24.2, pp. 118–125.

Collins, D L and A C Evans (1997). "Animal: Validation and Applications of Nonlinear Registration-Based Segmentation". In: *International Journal of Pattern Recognition and Artificial Intelligence* 11.08, pp. 1271–1294.

Corder, E H et al. (1993). "Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families." In: *Science* 261.5123, pp. 921–923.

Coupé, Pierrick et al. (2012). "Simultaneous segmentation and grading of anatomical structures for patient's classification: application to Alzheimer's disease." In: *NeuroImage* 59.4, pp. 3736–3747.

Crane, Paul K et al. (2012). "Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI)". In: *Brain Imaging and Behavior* 6.4, pp. 502–516.

Cummings, Jeffrey L, Travis Morstorf, and Kate Zhong (2014). "Alzheimer's disease drug-development pipeline: few candidates, frequent failures." In: *Alzheimer's research & therapy* 6.4, p. 37.

Dansereau, Christian et al. (2017). "A brain signature highly predictive of future progression to Alzheimer's dementia". In: *arXiv preprint arXiv:1712.08058*.

Davatzikos, Christos et al. (2011). "Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification." In: *Neurobiology of Aging* 32.12, 2322.e19–27.

Desikan, Rahul S et al. (2009). "Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease". In: *Brain* 132.8, pp. 2048–2057.

Devanand, D P et al. (2007). "Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease." In: *Neurology* 68.11, pp. 828–836.

Dick, Anthony Steven and Pascale Tremblay (2012). "Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language." In: *Brain* 135.Pt 12, pp. 3529–3550.

Dickerson, Bradford C et al. (2007). "Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment." In: *Archives of General Psychiatry* 64.12, pp. 1443–1450.

Dietterich, Thomas G (2000). "Ensemble methods in machine learning". In: *International workshop on multiple classifier systems*. Springer, pp. 1–15.

Dong, Aoyan et al. (2016). "CHIMERA: Clustering of Heterogeneous Disease Effects via Distribution Matching of Imaging Patterns". In: *Medical Imaging, IEEE Transactions on* 35.2, pp. 612–621.

Dong, Aoyan et al. (2017). "Heterogeneity of neuroanatomical patterns in prodromal Alzheimer's disease: links to cognition, progression and biomarkers." In: *Brain* 140.3, pp. 735–747.

Doyon, Julien et al. (2009). "Contributions of the basal ganglia and functionally related brain structures to motor learning." In: *Behavioural Brain Research* 199.1, pp. 61–75.

Dreiseitl, Stephan and Lucila Ohno-Machado (2002). "Logistic regression and artificial neural network classification models: a methodology review". In: *Journal of biomedical informatics* 35.5-6, pp. 352–359.

Dukart, Juergen and Alessandro Bertolino (2014). "When Structure Affects Function – The Need for Partial Volume Effect Correction in Functional and Resting State Magnetic Resonance Imaging Studies". In: *PLoS ONE* 9.12, e114227–18.

Eckerström, Carl et al. (2013). "A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia." In: *Journal of Alzheimer's disease : JAD* 36.3, pp. 421–431.

Edward, V et al. (2000). "Quantification of fMRI artifact reduction by a novel plaster cast head holder." In: *Human Brain Mapping* 11.3, pp. 207–213.

Elliott, M R, R W Bowtell, and P G Morris (1999). "The effect of scanner sound in visual, motor, and auditory functional MRI." In: *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 41.6, pp. 1230–1235.

Eskildsen, Simon F et al. (2013). "Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning." In: *NeuroImage* 65, pp. 511–521.

Fan, Rong-En et al. (2008). "LIBLINEAR: A library for large linear classification". In: *Journal of Machine Learning Research* 9.Aug, pp. 1871–1874.

Farrer, L A et al. (1997). "Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium." In: *JAMA : the journal of the American Medical Association* 278.16, pp. 1349–1356.

Ferreri, Florinda et al. (2011). "Motor cortex excitability in Alzheimer's disease: a transcranial magnetic stimulation follow-up study." In: *Neuroscience Letters* 492.2, pp. 94–98.

Fonov, Vladimir et al. (2011). "Unbiased average age-appropriate atlases for pediatric studies". In: *NeuroImage* 54.1, pp. 313–327.

Friedman, Lee and Gary H Glover (2006). "Report on a multicenter fMRI quality assurance protocol". In: *Journal of Magnetic Resonance Imaging* 23.6, pp. 827–839.

Friedman, Lee, Gary H Glover, and Fbirn Consortium (2006). "Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal-to-fluctuation-noise-ratio (SFNR) differences." In: *NeuroImage* 33.2, pp. 471–481.

Frisoni, Giovanni B et al. (2010). "The clinical use of structural MRI in Alzheimer disease". In: *Nature reviews. Neurology* 6.2, pp. 67–77.

Gibbons, Laura E et al. (2012). "A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment". In: *Brain Imaging and Behavior* 6.4, pp. 517–527.

Giove, Federico et al. (2009). "Images-based suppression of unwanted global signals in resting-state functional connectivity studies". In: *Magnetic Resonance Imaging* 27.8, pp. 1058–1064.

Gour, N et al. (2011). "Basal functional connectivity within the anterior temporal network is associated with performance on declarative memory tasks." In: *NeuroImage* 58.2, pp. 687–697.

Greicius, M D et al. (2004). "Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI." In: *Proceedings of the National Academy of Sciences of the United States of America* 101.13, pp. 4637–4642.

Guillozet, Angela L et al. (2003). "Neurofibrillary tangles, amyloid, and memory in aging and mild cognitive impairment." In: *Archives of Neurology* 60.5, pp. 729–736.

Haar, Shlomi et al. (2014). "Anatomical Abnormalities in Autism?" In: *Cerebral Cortex*.

Hardy, John and Dennis J Selkoe (2002). "The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics." In: *Science* 297.5580, pp. 353–356.

Hebert, Liesi E et al. (2013). "Alzheimer disease in the United States (2010-2050) estimated using the 2010 census." In: *Neurology* 80.19, pp. 1778–1783.

Hedden, Trey et al. (2009). "Disruption of functional connectivity in clinically normal older adults harboring amyloid burden." In: *Journal of Neuroscience* 29.40, pp. 12686–12694.

Humbert, Ianessa A et al. (2010). "Early deficits in cortical control of swallowing in Alzheimer's disease." In: *Journal of Alzheimer's disease : JAD* 19.4, pp. 1185–1197.

Hwang, Jihye et al. (2016). "Prediction of Alzheimer's disease pathophysiology based on cortical thickness patterns." In: *Alzheimers Dementia Diagnosis, Assessment Disease Monitoring* 2, pp. 58–67.

Iqbal, Khalid, Fei Liu, and Cheng-Xin Gong (2016). "Tau and neurodegenerative disease: the story so far". In: *Nature Reviews Neurology* 12.1, p. 15.

Irish, Muireann et al. (2011). "Everyday episodic memory in amnestic mild cognitive impairment: a preliminary investigation." In: *BMC neuroscience* 12.1, p. 80.

Jack, C R et al. (1997). "Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease." In: *Neurology* 49.3, pp. 786–794.

Jack, C R et al. (2013). "Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers". In: *The Lancet Neurology* 12.2, pp. 207–216.

Jain, Anil K, M Narasimha Murty, and Patrick J Flynn (1999). "Data clustering: a review". In: *ACM computing surveys (CSUR)* 31.3, pp. 264–323.

Jansen, Willemijn J et al. (2015). "Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis". In: *JAMA : the journal of the American Medical Association* 313.19, pp. 1924–1938.

Jones, David T et al. (2016). "Cascading network failure across the Alzheimer's disease spectrum." In: *Brain* 139.Pt 2, pp. 547–562.

Jones, Eric, Travis Oliphant, and Pearu Peterson (2001). "SciPy: open source scientific tools for Python, http://www.scipy.org/". In:

Jong, L W de et al. (2008). "Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study." In: *Brain* 131.Pt 12, pp. 3277–3285.

Kelly, Clare et al. (2012). "Characterizing variation in the functional connectome: promise and pitfalls". In: *Trends in Cognitive Sciences* 16.3, pp. 181–188.

Kluger, A et al. (1999). "Neuropsychological prediction of decline to dementia in nondemented elderly." In: *Journal of Geriatric Psychiatry and Neurology* 12.4, pp. 168–179.

Klunk, W E et al. (2007). "Amyloid deposition begins in the striatum of presenilin-1 mutation carriers from two unrelated pedigrees." In: *Journal of Neuroscience* 27.23, pp. 6174–6184.

Koch, Walter et al. (2012). "Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease". In: *Neurobiology of Aging* 33.3, pp. 466–478.

Korolev, Igor O et al. (2016). "Predicting Progression from Mild Cognitive Impairment to Alzheimer's Dementia Using Clinical, MRI, and Plasma Biomarkers via Probabilistic Pattern Classification". In: *PLoS ONE* 11.2, e0138866–25.

Liang, Peipeng et al. (2012). "Three subsystems of the inferior parietal cortex are differently affected in mild cognitive impairment." In: *Journal of Alzheimer's disease : JAD* 30.3, pp. 475–487.

Logothetis, Nikos K et al. (2001). "Neurophysiological investigation of the basis of the fMRI signal". In: *Nature* 412.6843, p. 150.

Lund, Torben E et al. (2006). "Non-white noise in fMRI: Does modelling have an impact?" In: *NeuroImage* 29.1, pp. 54–66.

Madsen, S K et al. (2010). "3D maps localize caudate nucleus atrophy in 400 Alzheimer's disease, mild cognitive impairment, and healthy elderly subjects." In: *Neurobiology of Aging* 31.8, pp. 1312–1325.

Martinez, M et al. (1998). "Apolipoprotein E epsilon4 allele and familial aggregation of Alzheimer disease." In: *Archives of Neurology* 55.6, pp. 810–816.

Maruyama, Masahiro et al. (2013). "Imaging of Tau Pathology in a Tauopathy Mouse Model and in Alzheimer Patients Compared to Normal Controls". In: *Neuron* 79.6, pp. 1094–1108.

Mathotaarachchi, Sulantha et al. (2017). "Identifying incipient dementia individuals using machine learning and amyloid imaging." In: *Neurobiology of Aging* 59, pp. 80–90.

McDonald, C R et al. (2009). "Regional rates of neocortical atrophy from normal aging to early Alzheimer disease." In: *Neurology* 73.6, pp. 457–465.

McKhann, Guy M et al. (2011). "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." In: *Alzheimer's & dementia : the journal of the Alzheimer's Association*. Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. guy.mckhann@jhu.edu. Elsevier, pp. 263–269.

Misra, Chandan, Yong Fan, and Christos Davatzikos (2009). "Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI". In: *NeuroImage* 44.4, pp. 1415–1422.

Mitchell, A J and M Shiri-Feshki (2009). "Rate of progression of mild cognitive impairment to dementia - meta-analysis of 41 robust inception cohort studies". In: *Acta Psychiatrica Scandinavica* 119.4, pp. 252–265.

Mitchell, Joanna et al. (2009). "Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm". In: *Journal of Neurology* 256.9, pp. 1500–1509.

Moradi, Elaheh et al. (2015). "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects". In: *NeuroImage* 104.C, pp. 398–412.

Mormino, Elizabeth C et al. (2011). "Relationships between β-amyloid and functional connectivity in different components of the default mode network in aging." In: *Cerebral Cortex* 21.10, pp. 2399–2407.

Murray, Melissa E et al. (2011). "Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study". In: *The Lancet Neurology* 10.9, pp. 785–796.

Noh, Young et al. (2014). "Anatomical heterogeneity of Alzheimer disease Based on cortical thickness on MRIs". In: *Neurology* 83.21, pp. 1936–1944.

Nutter-Upham, Katherine E et al. (2008). "Verbal fluency performance in amnestic MCI and older adults with cognitive complaints." In: *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists* 23.3, pp. 229–241.

Orban, Pierre et al. (2017). "Subtypes of functional brain connectivity as early markers of neurodegeneration in Alzheimer's disease". In: *bioRxiv*, p. 195164.

Ossenkoppele, Rik et al. (2015). "Prevalence of amyloid PET positivity in dementia syndromes: a meta-analysis". In: *JAMA : the journal of the American Medical Association* 313.19, pp. 1939–1949.

Packard, Mark G and Barbara J Knowlton (2002). "Learning and memory functions of the Basal Ganglia." In: *Annual review of neuroscience* 25, pp. 563–593.

Park, Hae-Jeong and Karl Friston (2013). "Structural and functional brain networks: from connections to cognition." In: *Science* 342.6158, pp. 1238411–1238411.

Park, Hyunjin et al. (2012). "Dimensionality reduced cortical features and their use in the classification of Alzheimer's disease and mild cognitive impairment". In: *Neuroscience Letters* 529.2, pp. 123–127.

Park, Jong-Yun et al. (2017). "Robust Identification of Alzheimer's Disease subtypes based on cortical atrophy patterns." In: *Scientific Reports* 7, p. 43270.

Pedregosa, Fabian et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research*, pp. 2825–2830.

Petersen, R C (2004). "Mild cognitive impairment as a diagnostic entity". In: *Journal of internal medicine* 256.3, pp. 183–194.

Petersen, R C et al. (2001). "Current concepts in mild cognitive impairment." In: *Archives of Neurology* 58.12, pp. 1985–1992.

Petersen, R C et al. (2014). "Mild cognitive impairment: a concept in evolution." In: *Journal of internal medicine* 275.3, pp. 214–228.

Petersen, Ronald C (2003). "Mild cognitive impairment clinical trials". In: *Nature Reviews Drug Discovery* 2.8, pp. 646–653.

Phipson, Belinda and Gordon K Smyth (2010). "Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn". In: *Statistical applications in genetics and molecular biology* 9.1.

Pievani, Michela et al. (2010). "Assessment of white matter tract damage in mild cognitive impairment and Alzheimer's disease". In: *Human Brain Mapping* 31.12, pp. 1862–1875.

Pinheiro, José et al. (2018). "nlme: linear and nonlinear mixed effects models. R package version 3.1-117, https://CRAN.R-project.org/package=nlme". In:

Power, Jonathan D et al. (2012). "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion". In: *NeuroImage* 59.3, pp. 2142–2154.

Qi, Zhigang et al. (2010). "Impairment and compensation coexist in amnestic MCI default mode network". In: *NeuroImage* 50.1, pp. 48–55.

Querbes, Olivier et al. (2009). "Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve". In: *Brain* 132.8, pp. 2036–2047.

Rathore, Saima et al. (2017). "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages". In: *NeuroImage* 155, pp. 530–548.

Rémy, Florence et al. (2015). "White matter disruption at the prodromal stage of Alzheimer's disease: Relationships with hippocampal atrophy and episodic memory performance". In: *NeuroImage. Clinical* 7.C, pp. 482–492.

Saad, Ziad S et al. (2012). "Trouble at rest: how correlation patterns and group differences become distorted after global signal regression." In: *Brain Connectivity* 2.1, pp. 25–32.

Samper-González, Jorge et al. (2018). "Reproducible evaluation of classification methods in Alzheimer's disease: framework and application to MRI and PET data". In: *bioRxiv*, p. 274324.

Seabold, Skipper and Josef Perktold (2010). "Statsmodels: Econometric and statistical modeling with python". In: *Proceedings of the 9th Python in Science Conference*. SciPy society Austin.

Seeley, W W et al. (2009). "Neurodegenerative Diseases Target Large-Scale Human Brain Networks". In: *Neuron* 62.1, pp. 42–52.

Shaw, Leslie M et al. (2009). "Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects". In: *Annals of Neurology* 65.4, pp. 403–413.

Shehzad, Zarrar et al. (2014). "A multivariate distance-based analytic framework for connectome-wide association studies." In: *NeuroImage* 93 Pt 1, pp. 74–94.

Sheline, Y I et al. (2010a). "Amyloid Plaques Disrupt Resting State Default Mode Network Connectivity in Cognitively Normal Elderly". In: *Biological Psychiatry* 67.6, pp. 584–587.

Sheline, Y I et al. (2010b). "APOE4 Allele Disrupts Resting State fMRI Connectivity in the Absence of Amyloid Plaques or Decreased CSF A 42". In: *Journal of Neuroscience* 30.50, pp. 17035–17040.

Small, Gary W et al. (2006). "PET of brain amyloid and tau in mild cognitive impairment." In: *The New England journal of medicine* 355.25, pp. 2652–2663.

Snyder, S W et al. (1994). "Amyloid-beta aggregation: selective inhibition of aggregation in mixtures of amyloid with different chain lengths". In: *Biophysical Journal* 67.3, pp. 1216–1228.

Sorg, Christian et al. (2007). "Selective changes of resting-state networks in individuals at risk for Alzheimer's disease." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.47, pp. 18760–18765.

Spillantini, Maria Grazia and Michel Goedert (2013). "Tau pathology and neurodegeneration". In: *The Lancet Neurology* 12.6, pp. 609–622.

Strimbu, Kyle and Jorge A Tavel (2010). "What are biomarkers?" In: *Current Opinion in HIV and AIDS* 5.6, p. 463.

Tabert, Matthias H et al. (2006). "Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment." In: *Archives of General Psychiatry* 63.8, pp. 916–924.

Tam, Angela et al. (2015). "Common Effects of Amnestic Mild Cognitive Impairment on Resting-State Connectivity Across Four Independent Studies". In: *Frontiers in Aging Neuroscience* 7.7, pp. 2266–14.

Van Dijk, K R A et al. (2010). "Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization." In: *Journal of Neurophysiology* 103.1, pp. 297–321.

Vanhoutte, G, M Verhoye, and A Van der Linden (2006). "Changing body temperature affects the T2* signal in the rat brain and reveals hypothalamic activity." In: *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 55.5, pp. 1006–1012.

Varol, Erdem et al. (2017). "HYDRA: Revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework". In: *NeuroImage* 145.Part B, pp. 346–364.

Villemagne, Victor L et al. (2009). "High striatal amyloid beta-peptide deposition across different autosomal Alzheimer disease mutation types." In: *Archives of Neurology* 66.12, pp. 1537–1544.

Villeneuve, Sylvia and William J Jagust (2015). "Imaging Vascular Disease and Amyloid in the Aging Brain: Implications for Treatment." In: *The journal of prevention of Alzheimer's disease* 2.1, pp. 64–70.

Visser, P-J, Philip Scheltens, and F R J Verhey (2005). "Do MCI criteria in drug trials accurately identify subjects with predementia Alzheimer's disease?" In: *Journal of neurology, neurosurgery, and psychiatry* 76.10, pp. 1348–1354.

Ward, Alex et al. (2013). "Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature." In: *Dementia and Geriatric Cognitive Disorders Extra* 3.1, pp. 320–332.

Wee, Chong-Yaw et al. (2013). "Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns." In: *Human Brain Mapping* 34.12, pp. 3411–3425.

Willer, Cristen J, Yun Li, and Gonçalo R Abecasis (2010). "METAL: fast and efficient meta-analysis of genomewide association scans." In: *Bioinformatics* 26.17, pp. 2190–2191.

Wu, Liyong et al. (2014). "Resting state executive control network adaptations in amnestic mild cognitive impairment." In: *Journal of Alzheimer's disease : JAD* 40.4, pp. 993–1004.

Yan, Chao-Gan et al. (2009). "Spontaneous Brain Activity in the Default Mode Network Is Sensitive to Different Resting-State Conditions with Limited Cognitive Load". In: *PLoS ONE* 4.5, e5743.

Yeo, B T Thomas et al. (2011). "The organization of the human cerebral cortex estimated by intrinsic functional connectivity". In: *Journal of Neurophysiology* 106.3, pp. 1125–1165.

Zhang, Hong-Ying et al. (2010). "Resting brain connectivity: changes during the progress of Alzheimer disease." In: *Radiology* 256.2, pp. 598–606.

Zhang, Xiuming et al. (2016). "Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease." In: *Proceedings of the National Academy of Sciences of the United States of America* 113.42, E6535–E6544.

Appendix A

Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis

AmanPreet Badhwar, Angela Tam, Christian Dansereau, Pierre Orban, Felix Hoffstaedter, Pierre Bellec

ELSEVIER

CrossMark

Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 8 (2017) 73-85

Alzheimer's & Dementia

Neuroimaging

Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis

AmanPreet Badhwar[a,b,*], Angela Tam[a,c,d], Christian Dansereau[a,b], Pierre Orban[a,b,d], Felix Hoffstaedter[e,f,g], Pierre Bellec[a,b,**]

[a]Centre de Recherche, Institut Universitaire de Gériatrie de Montréal, Montreal, Quebec, Canada
[b]Université de Montréal, Montreal, Quebec, Canada
[c]McGill University, Montreal, Quebec, Canada
[d]Douglas Mental Health University Institute Research Centre, Montreal, Quebec, Canada
[e]Institute of Neuroscience and Medicine (INM-1, INM-7), Research Centre Jülich, Jülich, Germany
[f]Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[g]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Abstract

Introduction: We performed a systematic review and meta-analysis of the Alzheimer's disease (AD) literature to examine consistency of functional connectivity alterations in AD dementia and mild cognitive impairment, using resting-state functional magnetic resonance imaging.
Methods: Studies were screened using a standardized procedure. Multiresolution statistics were performed to assess the spatial consistency of findings across studies.
Results: Thirty-four studies were included (1363 participants, average 40 per study). Consistent alterations in connectivity were found in the default mode, salience, and limbic networks in patients with AD dementia, mild cognitive impairment, or in both groups. We also identified a strong tendency in the literature toward specific examination of the default mode network.
Discussion: Convergent evidence across the literature supports the use of resting-state connectivity as a biomarker of AD. The locations of consistent alterations suggest that highly connected hub regions in the brain might be an early target of AD.
© 2017 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Resting-state fMRI; Functional connectivity; Alzheimer's disease; Mild cognitive impairment; Meta-analysis

1. Introduction

Alzheimer's disease (AD) exists on a continuum comprising a lengthy preclinical stage, a middle stage of mild cognitive impairment (MCI), and a final stage of dementia [1]. Symptoms usually start around the age of 65 years, except in rare patients with early onset (33–60 years) autosomal dominant AD (ADAD) [2,3]. Drugs currently available for AD provide limited short-term treatment of AD symptoms [4]. Trials of disease-modifying therapies for AD dementia patients have been unsuccessful, likely because intervention at this stage is too late to affect the neurodegenerative process. The focus now is on therapeutic intervention at the MCI and/or preclinical disease stages, with delay of dementia onset constituting a major clinical end point for clinical trials [1]. This approach depends on the identification of biomarkers that can aid early AD diagnosis [1,5]. Currently, validated AD biomarkers are (1) low cerebrospinal fluid (CSF) amyloid-β 42 levels and/or high

*Corresponding author. Tel.: +1-514-340-3540x3367; Fax: +1-514-340-2802.
**Corresponding author. Tel.: +1-514-340-3540x4782; Fax: +1-514-340-2802.
E-mail address: amanpreet.badhwar@criugm.qc.ca (A.B.), pierre.bellec@criugm.qc.ca (P.B.)

http://dx.doi.org/10.1016/j.dadm.2017.03.007
2352-8729/ © 2017 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

amyloid tracer retention on positron emission tomography (PET), indicating brain amyloidosis; (2) high CSF tau levels, indicating neuronal injury; (3) temporoparietal pattern of reduced 18F-fluorodeoxyglucose uptake on PET, indicating brain hypometabolism, and (4) patterns of brain atrophy on structural magnetic resonance imaging (MRI), indicating neurodegeneration [1,6].

Connectivity in resting-state functional magnetic resonance imaging (rsfMRI) is an emerging AD biomarker that holds promise for early diagnosis [1,5,7]. RsfMRI indirectly measures neural processing in the brain using blood oxygenation and can be used to identify spatially distributed networks [8]. The National Institute on Aging–Alzheimer's Association lists rsfMRI functional connectivity as a potential biomarker of neuronal injury, at an early stage of validation [6]. The existing literature is indeed mostly composed of proof-of-concept cross-sectional comparisons of cognitively healthy elderly individuals with patients suffering from mild (MCI) or severe (dementia) AD symptoms.

To date, multiple studies have reported intrinsic connectivity network (ICN) disturbances in patients with AD dementia and MCI, presymptomatic ADAD mutation carriers, and cognitively normal individuals carrying the at-risk APOEε4 allele and/or showing evidence of amyloidosis [9–12]. Despite such promising findings, the overall effect of AD on ICNs remains poorly characterized because of several inconsistencies in the literature, such as different acquisition protocols, processing methods, and/or exclusion/inclusion criteria [13]. Our aim was to perform a systematic review and meta-analysis to examine the consistency of intrinsic connectivity alterations in MCI and late-onset AD (LOAD) dementia across the literature. We also reviewed the burgeoning literature on connectivity abnormalities in ADAD and the at-risk APOEε4 genotype.

2. Methods

2.1. Literature search

We conducted a systematic review of PubMed articles up to December 3, 2015 in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [14]. Search terms and combinations used are provided in Supplementary Table 1. Results were filtered for duplicates within each of the two main search categories, that is, AD dementia or MCI patients (Fig. 1). Unique search results underwent further screening as described subsequently.

2.2. Study selection

Search results were subjected to two successive screenings with increasingly stringent criteria. The initial screen was performed on article abstracts. An article was included if the abstract indicated that it was a peer-reviewed original research article written in English and used rsfMRI to study LOAD and/or MCI in humans. Reviews, letters, case reports, and studies with subjects in whom MCI was associated with other diseases were omitted. Following the initial screening, we applied the following inclusion criteria: (1) used seed-based or independent component analysis rsfMRI methods; (2) investigated functional connectivity between patients (AD dementia or MCI) and age-matched healthy controls (HC); and (3) reported peak coordinates of significant statistical differences in average connectivity between groups and the direction of difference.

2.3. Data extraction

One reviewer (A.B.) conducted the searches and screened for duplicates. Two reviewers (A.B. and A.T.) independently screened all unique search results for potential inclusion in the meta-analysis. Only articles passing both reviewers' approval were considered for final inclusion. For each "included" article, coordinate data of significant between-group comparisons, such as AD versus HC, were transcribed by one reviewer and checked by two others (second reviewer [A.T.] and F.H.).

2.4. Meta-analysis

We performed complementary network- and voxel-based quantitative meta-analyses on six main group comparisons: pooled group with AD dementia and MCI patients termed ADMCI $<$ HC, ADMCI $>$ HC, MCI $<$ HC, MCI $>$ HC, AD $<$ HC, and AD $>$ HC. Although the voxel-based meta-analysis has finer spatial resolution for findings with high anatomic consistency, we assumed the network-based approach would have better sensitivity for detecting consistent involvement of anatomically distributed networks. Coordinates from articles using the same cohort were pooled under the PubMed unique identifier or PMID of the earliest publication and treated as results from a single study to avoid counting the cohort multiple times. Henceforth, an individual article will be referred to as a "study" and a group comparison yielding network and/or localization information (e.g., ADMCI $<$ HC) as a "contrast."

2.4.1. Network-based statistics

We performed network-based statistics on seed coordinates (seed statistics) to assess whether seed regions were preferentially selected from within certain networks in the literature. We also performed network-based statistics on coordinate data of significant contrasts (contrast statistics) to assess the consistency of network-level findings in the AD literature. In particular, we performed three types of contrast statistics: (1) all coordinates irrespective of seed network; and given the focus on the default mode network (DMN) in the literature, (2) coordinates associated with seeds inside the DMN only; and (3) coordinates associated with seeds outside the DMN, that is, non-DMN seeds. All analyses were conducted using a multiresolution atlas of group-level

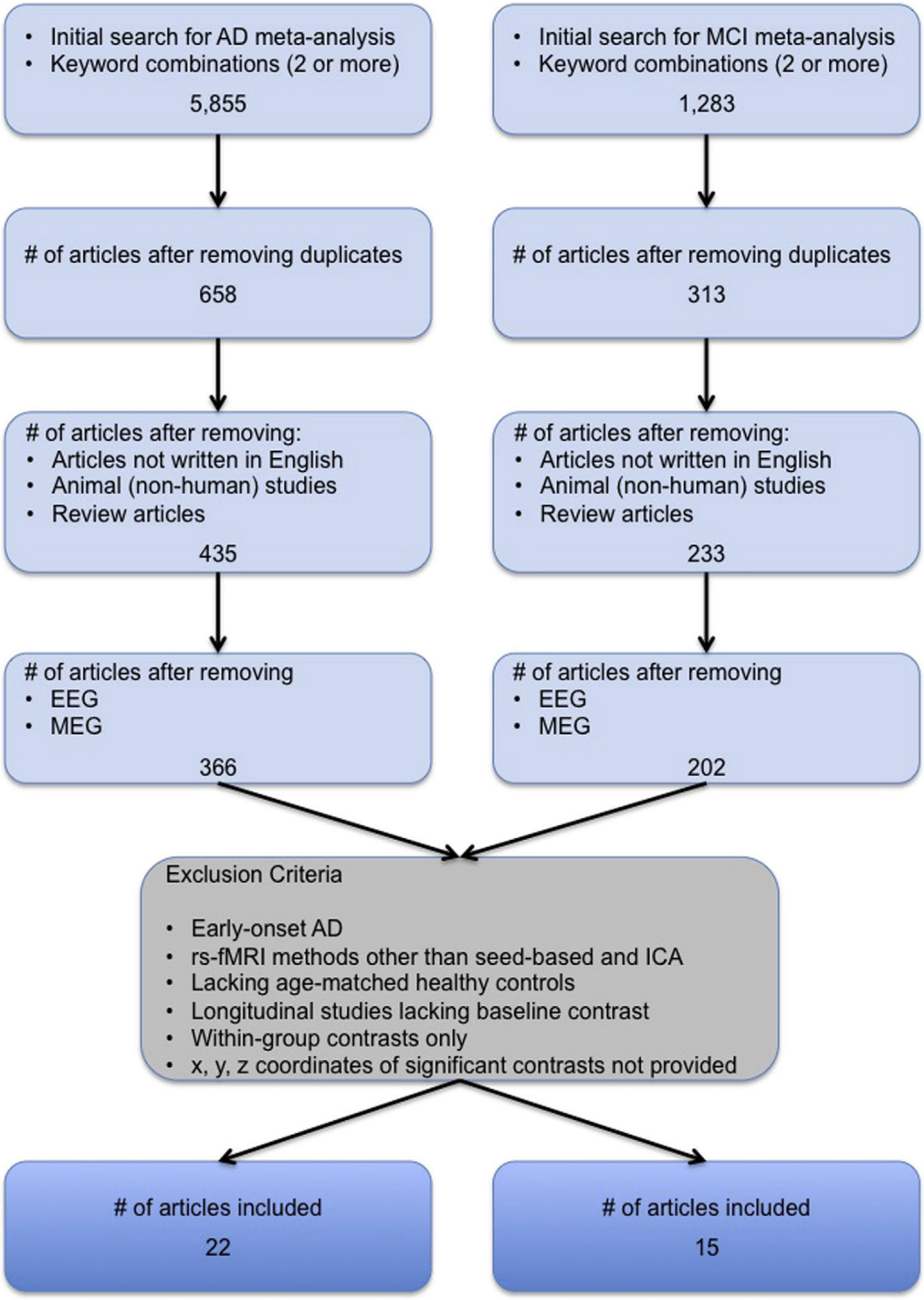

Fig. 1. Flowchart of the study selection process. Selection process for AD and MCI studies included in the meta-analyses. Studies using rsfMRI methods dissimilar to seed-based and ICA methods, such as degree centrality or graph theory, amplitude of low-frequency fluctuations, and regional homogeneity were not included. Abbreviations: AD, Alzheimer's disease; EEG, electroencephalogram; ICA, independent component analysis; MCI, mild cognitive impairment; MEG, magnetoencephalography; rsfMRI, resting-state functional magnetic resonance imaging.

functional brain parcellations derived from an independent rsfMRI data set, the Bootstrap Analysis of Stable Clusters–Cambridge atlas (https://dx.doi.org/10.6084/m9.figshare.1285615.v1) [15]. This atlas consists of nine functional parcellations capturing successively finer levels of spatial detail, of which we used parcellations at two resolutions: the first comprised seven commonly used large-scale networks (R7 atlas) and the second containing 36 networks (R36 atlas). We used R7 and R36 atlases for contrast statistics and only the R7 atlas for seed statistics. Because seeds were assigned indirectly for studies where coordinates were not provided, indirect assignment could not be performed with sufficient precision to use the R36 atlas. Assignment of seeds to one of the R7 networks was based on published coordinates,

when available. When only anatomic labels were provided for seed regions, network assignment was based on (1) the center of gravity in MNI space or (2) visual approximation if no further information was available. For independent component analysis-based studies, network assignment was based on (1) network coordinates when provided or (2) visual assignment to one or more of the seven networks based on the degree of spatial overlap.

We tested the spatial consistency of both seed and peak locations using the following approach. For each study, we computed the number of coordinates falling within each network, after conversion of Talairach space coordinates into MNI space using the Lancaster transform [16], when necessary. Coordinates falling outside of the gray matter mask (ICBM152) were assigned to the closest network. To remain unbiased to the number of coordinates reported per study, we computed the ratio of coordinates falling within each network to the total number of coordinates reported per study. This ratio was then averaged across studies. The significance of findings was assessed using Monte Carlo permutation tests. Using the total number of coordinates per study, we generated a random assignment of coordinates to networks, taking into consideration the volume of each network. Coordinate counts per network were normalized as described previously, followed by an averaging across studies. This Monte Carlo sampling process was repeated 10,000 times. Thereafter, we compared the distribution of the average frequency obtained from the random sampling with the frequency obtained from the meta-analysis, resulting in P value estimates [17]. Multiple comparisons across networks were accounted for using a false discovery rate (FDR) procedure (qFDR < 0.05) [18]. The P values less than .05 that did not survive multiple comparisons were deemed as "trends."

2.4.2. Voxel-based statistics

Voxel-level statistical analysis was performed using activation likelihood estimation (ALE), a widely used algorithm for coordinate-based meta-analysis of neuroimaging studies. ALE aims at delineating brain regions with above-chance convergence of reported coordinates across experiments [19]. Coordinates falling outside the gray matter mask were removed from the analysis. We used the in-house ALE algorithm implementation in MATLAB version 8.3.0.532, which treats each of the coordinates in a given experiment as a three-dimensional gaussian probability distribution centered at the given coordinate. The probability distributions acknowledge the spatial uncertainty associated with each experiment. For any given study, the width of the spatial uncertainty of its coordinates is determined based on empirical data on the between-subject and between-template variances representing the main components of this uncertainty [19]. Then, the probability distributions of all coordinates per included study are combined for each voxel, generating a modeled activation (MA) map. To limit the effect of multiple coordinates very close to one another within a given study, we used the "nonadditive" approach, which calculates MA maps by taking the maximum probability across overlapping gaussians [19]. ALE scores were computed on a voxel-by-voxel basis by taking the union across these MA maps. To distinguish between "true" and random convergence between studies (i.e., noise), ALE scores were compared with a null distribution reflecting a random spatial association between experiments (10,000 permutations). Nonparametric P values were assessed at a familywise error-corrected threshold of $P < .05$ on a cluster level (cluster-forming threshold: $P < .001$ at voxel level) and transformed into t scores for display purposes. Only contrasts including more than 18 experiments were considered, as recommended in a recent large-scale simulation study [20].

3. Results

3.1. Search results

The results of the initial search, along with studies systematically excluded from inclusion in our rsfMRI meta-analyses are presented in Fig. 1. Thirty-four studies totaling 1363 subjects (post pooling of identical cohorts) met our inclusion criteria and were included in the meta-analysis. The total included 352 MCI, 378 AD dementia (specifically LOAD), and 633 HC. Diagnostic criteria used per study for MCI and AD dementia are provided in the Supplementary Material (Supplementary Table 2 and Section 2). The bulk (54%) of the studies had 20 or less subjects per group. Twenty studies (66.7%) investigated rsfMRI connectivity measures with other domains, cognition being most frequent ($n = 11/22$ AD studies, $n = 9/15$ MCI studies), and few with levels of amyloid burden using Pittsburgh compound B ($n = 3$), brain atrophy ($n = 3$), and structural connectivity ($n = 1$). Alterations in functional connectivity were often ($n = 5/9$ studies) reported to be significantly correlated with episodic verbal learning and memory in MCI cohorts. Table 1 provides additional characteristics of the included rsfMRI studies, including scanner make, model, and strength, and seed region and/or ICN investigated. A summary of commonly used preprocessing steps utilized by the studies present in our meta-analysis are provided in Supplementary Table 3.

3.2. Network-based meta-analysis

3.2.1. Seed statistics

Using network-level statistics, we demonstrated that a disproportionately large number of studies specifically targeted the DMN (Fig. 2) irrespective of the population (ADMCI, MCI, or AD dementia) being studied.

3.2.2. Contrast statistics

We first examined R7 network-level statistics and all seeds combined. Aberrant functional brain connectivity was observed in ADMCI, MCI, and AD, relative to HC (Fig. 3). In the ADMCI cohort, we found both significant

Table 1
Characteristics of rsfMRI studies included in the meta-analysis

		AD					HC					AD				
Study	*N*	*n*	M	F	Age	SD	*n*	M	F	Age	SD	<HC	>HC	Scanner	Method	Seed region/ICN investigated
Wang et al. [63]	28	14	7	7	70.2	6.3	14	7	7	69.6	5.5	x		1.5 T S	SB	PCC
Zhang et al. [64][a]	32	16	6	10	71.6	5.1	16	7	9	71.3	4.9	x	x	1.5 T P	SB	PCC
Zhang et al. [65][a]	55	39	18	21	73.4		16	7	9	71.3	4.9	x	x	1.5 T P	SB	PCC
Sheline et al. [66]	83	35					48					x	x	3.0 T S	SB	Precuneus
Zhou et al. [52]	24	12	5	7	63.3	7.7	12	5	7	62.0		x	x	1.5/3.0/4.0 T S/GE/B	*SB* and ICA	AG (l), pregenual ACC (r)
Gili et al. [67]*	21	11	7	4	71.9	7.9	10	7	3	64.1	10.5	x		3.0 T S	*SB* and ICA	PCC, mPFC
Wu et al. [68][b]	31	15	6	9	64.0	8.3	16	7	9	65.0	9.2	x		3.0 T S	ICA	DMN
Li et al. [69][b]	31	15	6	9	64.0	8.3	16	7	9	65.0	9.2	x		3.0 T S	ICA	ATN (d, v)
Damoiseaux et al. [70]	39	21	9	12	64.2	8.7	18	12	6	62.7	10.3	x	x	3.0 T GE	ICA	DMN (a, p, v), SMN
Binnewijzend et al. [71]†	82	39	23	16	67.0	8.0	43	23	20	69.0	7.0	x		1.5 T S	ICA	DMN, working memory (l, r), visuospatial attention (d), spatial attention (v), SMN, auditory language, prVIS, sVIS, basal ganglia cerebellum
Kenny et al. [72]	32	16			77.3	8.9	16			76.3	8.3		x	3.0 T P	SB	Hippocampus (l, r), PCC, precuneus, prVIS
Zhu et al. [73]†	22	10	7	3	72.9	7.9	12	5	7	73.8	6.5	x		3.0 T GE	SB	ICC (l, r)
Balthazar et al. [74]	37	20			73.9	8.2	17			72.3	6.4	x	x	3.0 T P	ICA	DMN (d, v), SN (a, p)
Yao et al. [75][c,†]	62	35	12	23	72.4	8.5	27	16	11	69.2	6.5	x		3.0 T GE	SB	Amygdala (l, r)
Zhou et al. [76][c,†]	62	35	12	23	72.4	8.5	27	16	11	69.2	6.5	x	x	3.0 T GE	SB	T
Zhang et al. [77][c,†]	62	35	12	23	72.4	8.5	27	16	11	69.2	6.5	x		3.0 T GE	SB	MrD (l, r)
Gour et al. [29]	28	14	6	8	75.1	2.9	14	4	10	72.8	3.0	x	x	3.0 T S	SB	PCC, perirhinal cortex (l, r), dlPFC (l, r)
Weiler et al. [78]	48	22	6	16	73.4	5.7	26	6	20	70.0	6.6	x	x	3.0 T P	SB	PCC, Wernicke's (l); Broca's (l), dlPFC (l, r), saVC
Balachandar et al. [79]	30	15	9	6	67.3	6.6	15	9	6	64.4	8.9	x	x	3.0 T S	ICA	DMN, thalamic, ECN
Pasquini et al. [80]*	43	21	8	13	72.3	8.6	22	6	16	66.3	9.0	x	x	3.0 T P	ICA	DMN (a, p)
Adriaanse et al. [28]	59	28	17	11	72.0	4.9	31	17	14	72.0	4.3	x		1.5 T S	ICA	DMN, VIS (med, lat), AN, SMN, ECN, dorsovisual (l, r)
Yi et al. [81]*	23	11	1	10	64.2	2.4	12	3	9	71.8	1.2	x		3.0 T GE	ICA	DMN, SN

(*Continued*)

Table 1
Characteristics of rsfMRI studies included in the meta-analysis (*Continued*)

		MCI					HC					MCI				
Study	*N*	*n*	M	F	Age	SD	*n*	M	F	Age	SD	<HC	>HC	Scanner	Method	Seed region/ICN investigated
Sorg et al. [82]	40	24	13	11	69.3	8.1	16	10	6	68.1	3.8	x		1.5 T S	ICA	VIS, AN, ATN (v), spatial attention, DMN
Bai et al. [83]	56	30	15	15	72.5	4.4	26	12	14	71.6	5.3	x	x	1.5 T GE	SB	PCC
Gili et al. [67]*	20	10	6	4	71.2	4.1	10	7	3	64.1	10.5	x		3.0 T S	*SB* and ICA	PCC, mPFC
Bai et al. [84]	44	26	19	7	71.4	4.3	18	10	8	70.3	4.7		x	1.5 T GE	ICA	DMN
Xie et al. [85]	56	30	19	11	72.6	4.8	26	14	12	70.3	4.8	x		1.5 T GE	SB	Postcentral gyrus (l), hippocampus (l), medialFC (l), middleFC (l), precuneus (l, r), insula (l, r)
Jin et al. [86]	16	8	5	3	60.6	3.2	8	4	4	60.9	8.3	x	x	3.0 T GE	ICA	DMN
Han et al. [87]	80	40	7	33	86.3	4.5	40	15	25	86.3	4.5	x	x	1.5 T GE	SB	PCC
Liang et al. [88]	32	16	6	10	68.5	7.8	16	6	10	67.2	8.4	x	x	3.0 T S	SB	AG (l, r), supramarginal gyrus (l, r), intraparietal sulcus (r)
Hahn et al. [89]†	54	28	14	14	69.5	7.1	26	10	16	65.5	7.8	x		3.0 T P	ICA	DMN (a, p), ATN (d, v), ECN (l, r), SMN, VIS
Myers et al. [90]	35	23	14	9	69.3	7.4	12	5	9	63.8	5.2	x		3.0 T P	ICA	DMN (a, p), ATN (l, r, d), SN, prAN
Koch et al. [91]	40	24	14	10	68.2	8.4	16	7	9	64.8	5.4	x		3.0 T P	ICA	DMN (a, p), ATN (l, r, d), SN, prAN
Pasquini et al. [80]*	44	22	11	11	65.3	8.7	22	6	16	66.3	9.0		x	3.0 T P	ICA	DMN (a, p)
Das et al. [92]	69	30	14	16	71.6	6.8	39	18	21	70.6	9.0	x		3.0 T S	SB	Hippocampal subregions
Gardini et al. [93]	42	21	13	8	70.6	4.7	21	7	14	69.8	6.5		x	3.0 T GE	SB	PCC, mPFC
Yi et al. [81]*	32	20	4	16	71.0		12	3	9	71.8	1.2	x	x	3.0 T GE	ICA	DMN, SN

Abbreviations: a, anterior; ACC, anterior cingulate cortex; AD, Alzheimer's disease; AG, angular gyrus; AN, auditory network; ATN, attentional network; B, Brucker; d, dorsal; dlPFC, dorsolateral prefrontal cortex; DMN, default mode network; ECN, executive control network; F, female; GE, General Electrics; HC, healthy control; ICA, independent component analysis; ICC, isthmus of cingulate cortex; ICN, intrinsic connectivity network; l, left; lat, lateral; M, male; MCI, mild cognitive impairment; med, medial; medialFC, medial frontal cortex; middleFC, middle frontal cortex; mPFC, medial prefrontal cortex; MrD, marginal division; n, number of subjects; p, posterior; P, Philips; PCC, posterior cingulate cortex; prAN, primary auditory network; prVIS, primary visual network; r, right; rsfMRI, resting-state functional magnetic resonance imaging; S, Siemens; saVC, secondary associative visual cortex; SB, seed based; SMN, sensorimotor network; SD, standard deviation; sVIS, secondary visual network; T, Tesla; v, ventral.

NOTE. Data provided in "bold" indicate seven studies using shared cohorts. Coordinates from these seven studies were subsequently pooled under four studies (indicated by superscript letters a, b, and c), under the corresponding earliest publication using the cohort. In column "Method", when both seed-based and ICA rsfMRI methods were used by a study, the method given in "italics" indicates the method associated with reported coordinates. For column "Seed region/ICN investigated", all seed regions and ICNs investigated are listed, irrespective of significant findings.

*Studies reporting significant coordinates for both AD and MCI patients, relative to matched HC.

†Studies investigating both AD dementia and MCI cohorts.

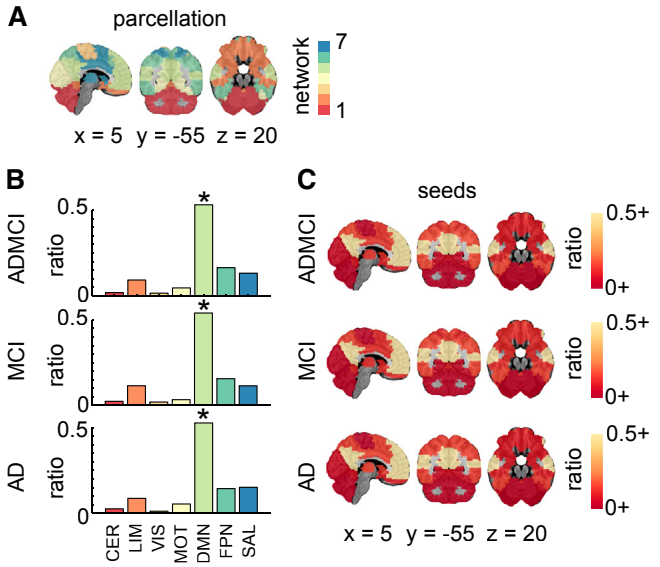

Fig. 2. Seed region network-level findings. (A) R7 atlas; (B) histograms showing the ratio of counts (or hits) across the seven networks for all seeds in ADMCI, MCI, and AD. Significant (* denoting qFDR < 0.05) prevalence of seeds in the DMN was demonstrated across all three cohorts; (C) seed region hit maps (ratio of dysconnectivity coordinates in each network) at R7. Maps are superimposed onto the anatomic International Consortium for Brain Mapping (ICBM) 152 template. x, y, and z Montreal Neurological Institute (MNI) coordinates are given for sagittal, coronal, and axial slices. Abbreviations: AD, Alzheimer's disease; ADMCI, AD dementia and MCI; CER, cerebellar network; DMN, default mode network; FDR, false discovery rate; FPN, frontoparietal network; HC, healthy control; LIM, limbic network; MCI, mild cognitive impairment; MOT, motor network; SAL, salience network; VIS, visual network.

hypoconnectivity and hyperconnectivity in the DMN. Significant hyperconnectivity in the DMN and limbic network (LIM) was observed in the MCI cohort. There was also significant hypoconnectivity in the DMN for the AD group, which appeared as a trend for the MCI group.

We then refined the spatial localization of effects found in R7 using the R36 atlas. Significant DMN hypoconnectivity in AD and ADMCI cohorts was detected in the precuneus (PCu) and posterior cingulate cortex (PCC) (Fig. 4). A trend for DMN hyperconnectivity was observed in the PCu for ADMCI and in both the PCu and PCC in MCI (Fig. 4). The LIM hyperconnectivity was observed as a trend in the hippocampus and entorhinal cortex in MCI patients (Fig. 4).

Finally, we investigated the robustness of findings with respect to the selection of seeds (DMN, non-DMN, or all combined), using the R7 atlas. Significant network-level findings derived from all seeds combined, as reported previously, replicated when using DMN seeds alone (Fig. 3A). In addition, a trend toward hypoconnectivity in MCI became significant using DMN seeds only. When focusing on non-DMN seed studies, no significant effects were observed in the DMN, as expected. The only significant result was hyperconnectivity of the salience network (SAL) in ADMCI, also present as a trend in AD subjects.

3.3. Voxel-based meta-analysis

ALE results demonstrated significant hypoconnectivity in the PCC and PCu in the ADMCI and AD studies (Fig. 5, Supplementary Table 4), consistent with our network-level findings using R7 and R36 atlases. This observation was made both for all seeds combined and DMN-only seeds (Fig. 5, Supplementary Table 4).

Unlike the network-level analysis, using ALE we found diminished connectivity in the primary visual cortex, both in ADMCI and AD. This was observed for all seeds combined as well as for DMN-only seeds in ADMCI and DMN-only seeds in AD. Finally, significant hyperconnectivity was observed in AD in the anterior insula (Fig. 5, Supplementary Table 4), consistent with the trend in the LIM observed using the R36 atlas.

4. Discussion

We report on a systematic meta-analysis of rsfMRI brain connectivity dysfunction in LOAD, using voxel-, region-, and network-level statistics. Our results demonstrated consistent connectivity alterations both within and outside of the DMN.

4.1. Connectivity changes in the DMN

4.1.1. Late-onset AD

Our results revealed a consistent decrease in DMN connectivity in the ADMCI and AD cohorts, particularly in the PCu and PCC, for all resolutions of meta-analysis. This finding is in line with previous meta-analyses centered on the DMN [21,22], and a recent study published after we completed our analysis [23]. DMN deterioration appears robust to the choice of analytical approaches, as previous meta-analyses largely included studies measuring regional homogeneity and amplitude of low-frequency fluctuation. Moreover, our results support previous literature reporting on the vulnerability of the DMN to multiple AD pathophysiology [24].

Unlike our robust findings in AD subjects, DMN hypoconnectivity in MCI could only be demonstrated using network-level statistics, suggesting a weaker, more distributed effect in MCI. However, we recently reported decreased DMN connectivity in a large multisite MCI cohort with a connectome-wide approach [13]. The modest findings of our present meta-analysis may be because of a lack of statistical power from having multiple, small, single-site samples. Clinical heterogeneity might also have played a role, that is, only a subset of MCI patients develop AD dementia [6,25], and there may be pathologic subtypes [26]. We also demonstrated DMN hyperconnectivity in MCI and ADMCI using network-level statistics. These changes may reflect both functional disconnection and compensation in response to damage at earlier stages of neurodegeneration, as well as direct or indirect pathologic mechanisms [27]. Moreover,

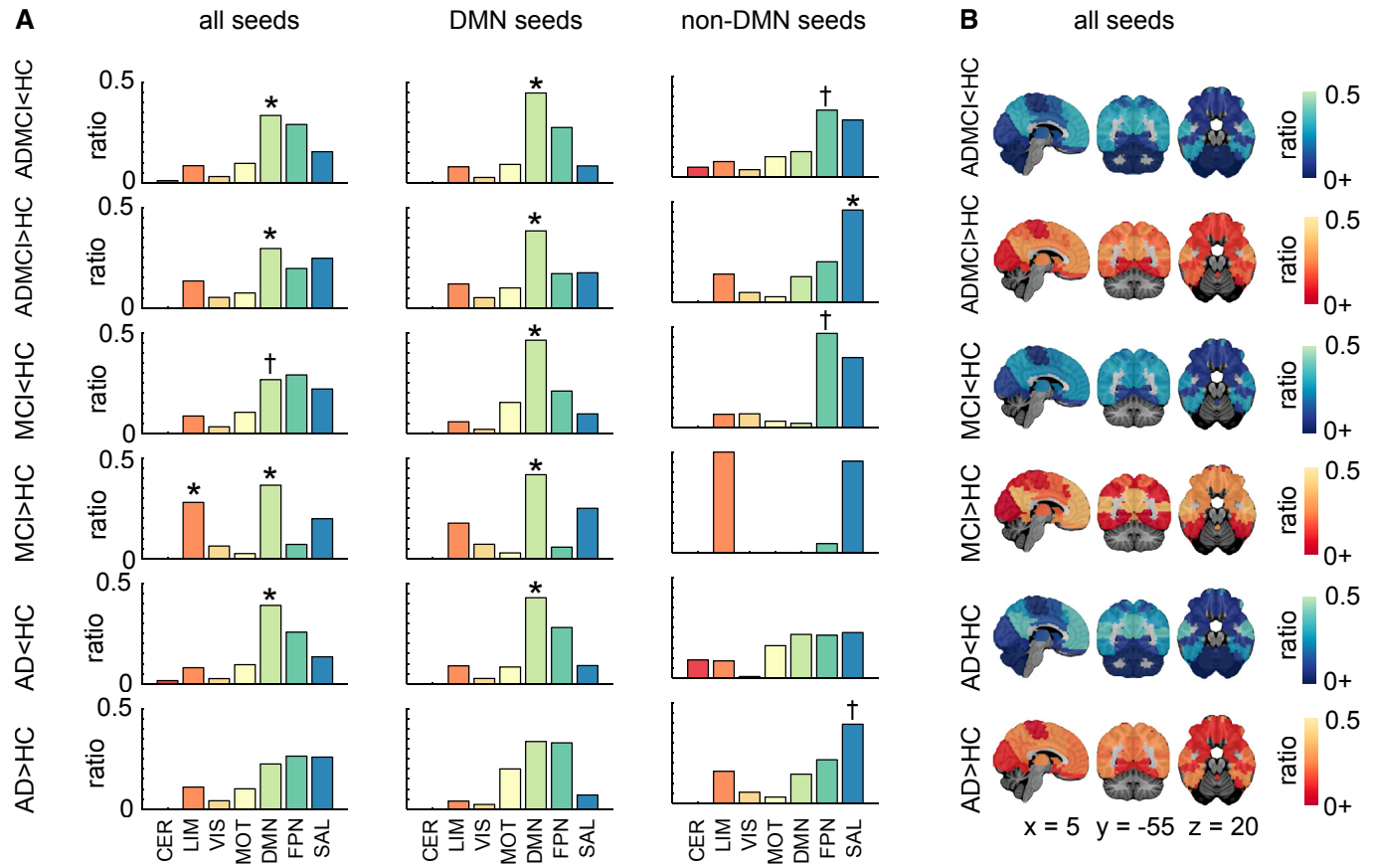

Fig. 3. Network-level findings using the R7 atlas. (A) Histograms showing per contrast the ratio of hits across the seven networks for all seeds, DMN seeds only, and non-DMN seeds. Networks with significant count (or hit) ratios are indicated by * denoting qFDR < 0.05, whereas † denotes $P < .05$ uncorrected. (B) Hit maps at R7 are shown for contrasts ADMCI < HC, ADMCI > HC, MCI < HC, MCI > HC, AD < HC, and AD > HC. Maps are superimposed onto the anatomic ICBM 152 template. x, y, and z MNI coordinates are given for sagittal, coronal, and axial slices. Abbreviations: AD, Alzheimer's disease dementia; ADMCI, AD dementia and MCI; CER, cerebellar network; DMN, default mode network; FDR, false discovery rate; FPN, frontoparietal network; HC, healthy control; LIM, limbic network; MCI, mild cognitive impairment; MOT, motor network; SAL, salience network; VIS, visual network.

there is some uncertainty of the specific nodes that actually show aberrant connectivity in our network-level analysis. This may give rise to apparent contradictory results.

4.1.2. Early onset AD

DMN hypoconnectivity of similar magnitude to LOAD was demonstrated in early onset non–ADAD [28,29], whereas in ADAD, DMN hypoconnectivity was slightly more pronounced than that in LOAD [30]. Altered DMN connectivity was observed in asymptomatic mutation carriers (*PSEN1*, *PSEN2*, or *APP*) many years before the age at which they were expected to develop symptoms [31–33], suggesting that aberrant connectivity may be a very early biomarker for AD.

4.1.3. Cognitively normal individuals at genetic risk for LOAD

Altered DMN connectivity has been reported in cognitively normal APOEε4 carriers compared with non-APOEε4 carriers. These alterations were found across all age groups, that is, elderly [12,34–36], middle-aged [37–39], and young adults [40,41], and were associated with worse cognition in middle-aged and elderly carriers [35,37,39]. Studies have also reported connectivity changes in the DMN in the absence of Pittsburgh compound B–detectable brain amyloidosis [12,40,41], further validating the potential of rsfMRI connectivity as an early marker of synaptic and neuronal dysfunction in AD.

4.1.4. Cognitively normal elderly at risk for LOAD

Aberrant DMN dysconnectivity, particularly reduced connectivity between the anterior and posterior DMN, has been associated with aging and age-related cognitive decline [33,42]. DMN hypoconnectivity may arise as early as middle age [43,44], with decreases occurring at differing rates between sexes [45] most likely due to the differential effect of sex on AD risk [46]. Reduced DMN integrity has also been reported in cognitively normal elderly with abnormal levels of CSF amyloid or tau proteins [47], as well as PET-detectable cerebral amyloidosis [48]. These results suggest that some of the effects related to normal aging in the literature may be driven by preclinical AD. Very few

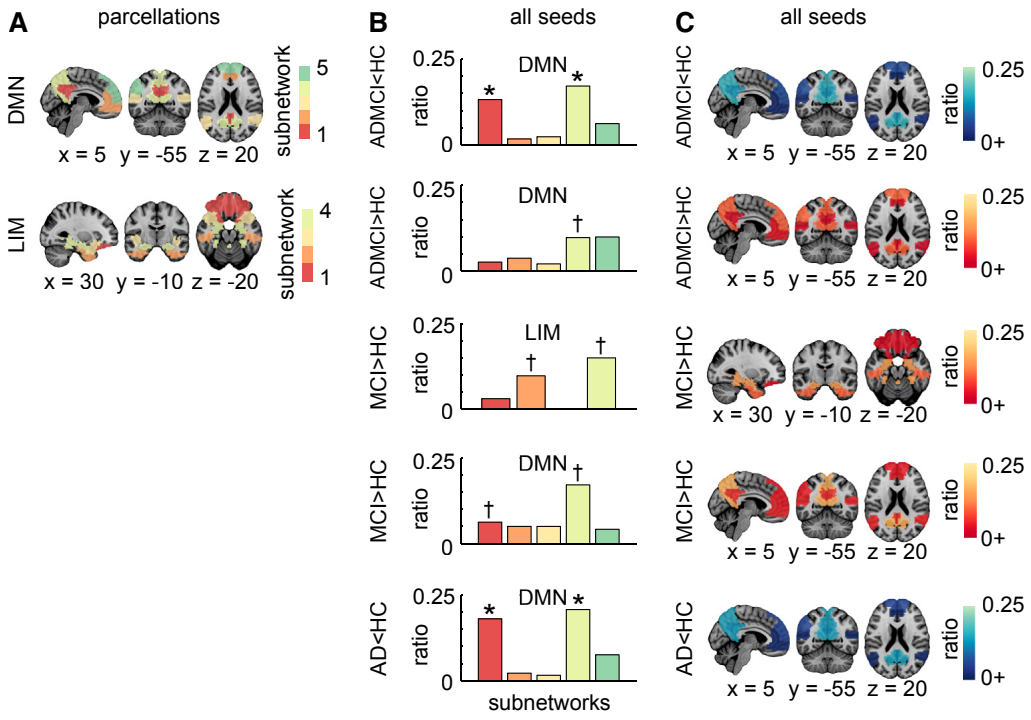

Fig. 4. Network-level findings using the R36 atlas. (A) Functional template at R36 showing the breakdown of the DMN and LIM into subnetworks. These two networks were significant (qFDR < 0.05 for contrasts ADMCI < HC for DMN, ADMCI > HC for DMN, MCI > HC for DMN and LIM, AD < HC for DMN) or trended toward significance ($P < .05$ uncorrected for contrasts MCI < HC for DMN) for the "all-seeds" condition at R7. (B) Histograms showing per selected contrast (as described in A), the ratio of counts (or hits) across the subnetworks. Subnetworks with significant hit ratios are indicated by * denoting qFDR < 0.05, whereas † denotes a trend with $P < .05$ uncorrected. (C) Hit maps at R36 for brain regions that overlap with significant or trending toward significance networks (as described in A). Maps are superimposed onto the anatomic ICBM 152 template. x, y, and z MNI coordinates are given for sagittal, coronal, and axial slices. Abbreviations: AD, Alzheimer's disease dementia; ADMCI, AD dementia and MCI; CER, cerebellar network; DMN, default mode network; FDR, false discovery rate; FPN, frontoparietal network; HC, healthy control; LIM, limbic network; MCI, mild cognitive impairment; MOT, motor network; SAL, salience network; VIS, visual network.

studies examined the interactions between age, sex, LOAD, and rsfMRI connectivity, which is clearly an important avenue for future work.

4.2. Connectivity changes outside the DMN

Our meta-analysis confirmed that intrinsic connectivity disruptions in LOAD are not confined to the DMN. We found increased connectivity in the SAL in ADMCI and AD. Abnormal SAL connectivity has now been reported in another LOAD study [49] published after we completed our meta-analysis and has also been demonstrated in ADAD [30], APOEε4 carriers [36,37], and the elderly [50], with connectivity increases highlighted in APOEε4 carriers. With the anterior insula as a key hub, the SAL plays a pivotal role in network switching between the DMN and frontoparietal network (FPN), two networks exhibiting competitive interactions during cognitive information processing [51]. Association of heightened SAL connectivity with reduced DMN connectivity in AD suggests that progressive DMN impairment may be deleterious to SAL function [52].

We also found increased connectivity in the LIM in MCI. Heightened LIM connectivity has been reported in early onset, non-ADAD patients [29], and in individuals with subjective memory impairment [53]. The effect of APOEε4 carriage on LIM connectivity, however, lacks consensus [54–56]. Since LIM hyperconnectivity in early onset AD patients was shown to correlate positively with memory performance, it is likely that increased connectivity in this network contributes to preserving function in the face of medial temporal lobe pathology [29].

4.3. Selective vulnerability of multimodal networks in AD

The DMN, SAL, and FPN are multimodal networks that interconnect cortical regions associated with various cognitive functions, and they have been demonstrated computationally to support integrative information processing at the cost of being vulnerable to early and fast spreading of insults [57]. Supporting this theoretical finding is the recent observation that tau and amyloid-β, despite their independent patterns of spatial deposition, overlap with brain tissue loss in hub regions of multimodal networks [58]. These multimodal networks are also metabolically expensive and display higher rates of cerebral blood flow, aerobic glycolysis, and oxidative

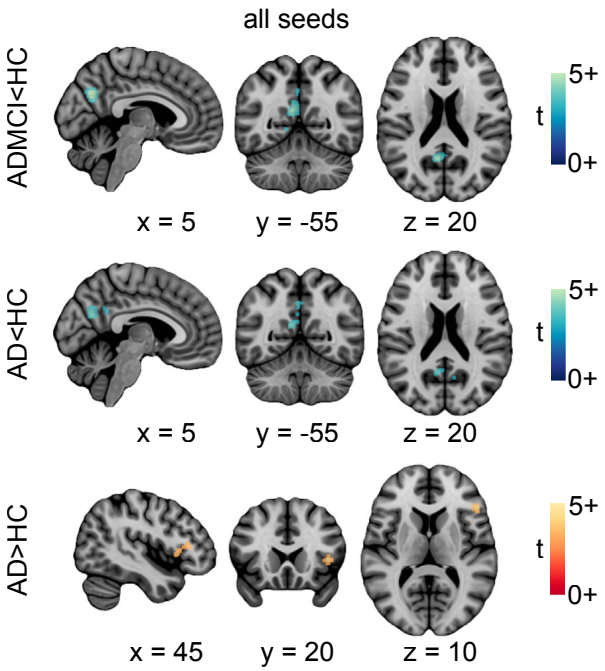

Fig. 5. Location of significant convergence of the voxel-level findings. Regions exhibiting significant rsfMRI abnormalities for contrasts ADMCI < HC, AD < HC, and AD > HC. Activation likelihood estimation images were thresholded at $P < .05$ (cluster-level family wise error or FWE corrected for multiple comparisons; cluster-forming threshold $P < .001$ at the voxel level) and displayed as t scores, with hyperconnectivity in red-orange and hypoconnectivity in blue-green. Maps are superimposed onto the anatomic ICBM 152 template. x, y, and z MNI coordinates are given for sagittal, coronal, and axial slices. Abbreviations: AD, Alzheimer's disease dementia; ADMCI, AD dementia and mild cognitive impairment; HC, healthy control; rsfMRI, resting-state functional magnetic resonance imaging.

glucose metabolism [59]. The high-value/high-cost characteristics of the DMN, SAL, and FPN may make them vulnerable to AD-associated pathogenic processes, such as metabolic dysfunction/oxidative stress, and accumulation of toxic proteins, such as amyloid-β [59]. The hypothesis that multimodal networks/regions are particularly susceptible to AD-associated pathophysiological processes may explain our finding of consistent alterations of these networks.

4.4. Limitations

Our literature search did not identify an abundance of rsfMRI literature in AD and MCI cohorts, which clearly expresses the need for additional research. The relatively low number of experiments that met our inclusion criteria might have underpowered our voxel-level findings, especially for the MCI contrasts. In addition, our search demonstrated that typical studies featured small samples, and also that analytical methods were quite variable in the field (a main reason for excluding an article was due to methodology used). This setting is particularly amenable to questionable research practices, including "p-hacking" (testing several methods, reporting only one). Given the near absence of negative results reporting in the field, on one hand, and the large size of the rsfMRI field, on the other hand, there is no question that some amount of publication bias is also present. Meta-analytical tools, such as funnel plots, are available to detect both selective reporting and p-hacking but are not feasible given current reporting practices in the rsfMRI community [60].

Another limitation of our study is experimental heterogeneity, in terms of population recruitment, scan acquisition (e.g., scanner make and model, scanning parameters), and processing choices [13,20,61]. The prominence of the DMN in our results partly reflects the focus on this network in the literature, which we quantified using seed statistics. Hypothesis-driven analyses on the DMN are attractive for assessing connectivity changes in small samples; as such analyses will have good statistical power if the DMN truly carries the larger effects in the brain. However, full-brain studies will be required to get a more comprehensive view on AD-related changes in rsfMRI network connectivity using meta-analyses. The current trend toward large public samples [13,62] is enabling unbiased meta-analyses, pooling neuroimaging data across many studies instead of relying on published coordinates. This will hopefully resolve most of the aforementioned limitations in the future.

5. Conclusions

Our meta-analysis demonstrated consistent connectivity alterations in the DMN, SAL, and LIM in the spectrum of LOAD, supporting the use of resting-state connectivity as a biomarker of AD.

Acknowledgments

The computational resources used to perform the data analysis were provided by Compute Canada (www.computecanada.org) and CLUMEQ (www.clumeq.mcgill.ca), which is funded in part by NSERC (MRS), FQRNT, and McGill University.
This research was supported by the Canadian Consortium on Neurodegeneration in Aging. The Canadian Consortium on Neurodegeneration in Aging is supported by a grant from the Canadian Institutes of Health Research with funding from several partners including the Alzheimer Society of Canada, Sanofi, and Women's Brain Health Initiative. This research was also supported by the Courtois Foundation (P.B.) and an Alzheimer Society Postdoctoral Fellowship (A.B.).

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.dadm.2017.03.007.

RESEARCH IN CONTEXT

1. Systematic review: We conducted a systematic review of PubMed-indexed resting-state functional magnetic resonance imaging (rsfMRI) studies in accordance with the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" guidelines. We included studies that investigated differences in functional connectivity, relative to controls, between patients with Alzheimer's disease (AD) and/or mild cognitive impairment, and reported coordinates of findings.
2. Interpretation: Typical rsfMRI functional connectivity studies in AD suffer from low statistical power. Our meta-analysis quantifies if and where convergent findings have been reported in the literature and strengthens the evidence for the use of rsfMRI as an AD biomarker.
3. Future directions: A disproportionately large portion of studies specifically investigated the default mode network, based on well-grounded hypotheses on AD pathophysiology. It is unclear if AD truly has larger effects on default-mode connectivity because of limited power to examine other networks. Future research should aim for full-brain investigations using larger study populations.

References

[1] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011; 7:280–92.
[2] Wu L, Rosa-Neto P, Hsiung GY, Sadovnick AD, Masellis M, Black SE, et al. Early-onset familial Alzheimer's disease (EOFAD). Can J Neurol Sci 2012;39:436–45.
[3] Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, et al. Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. Am J Hum Genet 1999;65:664–70.
[4] Hughes RE, Nikolic K, Ramsay RR. One for all? Hitting multiple Alzheimer's disease targets with one drug. Front Neurosci 2016;10:177.
[5] Matthews PM, Hampshire A. Clinical concepts emerging from fMRI functional connectomics. Neuron 2016;91:511–28.
[6] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011; 7:270–9.
[7] Vemuri P, Jones DT, Jack CR Jr. Resting state functional MRI in Alzheimer's disease. Alzheimers Res Ther 2012;4:2.
[8] Logothetis NK. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. Philos Trans R Soc Lond B Biol Sci 2002;357:1003–37.
[9] Krajcovicova L, Marecek R, Mikl M, Rektorova I. Disruption of resting functional connectivity in Alzheimer's patients and at-risk subjects. Curr Neurol Neurosci Rep 2014;14:491.
[10] Brier MR, Thomas JB, Ances BM. Network dysfunction in Alzheimer's disease: refining the disconnection hypothesis. Brain Connect 2014;4:299–311.
[11] Jack CR Jr, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 2013;12:207–16.
[12] Sheline YI, Morris JC, Snyder AZ, Price JL, Yan Z, D'Angelo G, et al. APOE4 allele disrupts resting state fMRI connectivity in the absence of amyloid plaques or decreased CSF Aβ42. J Neurosci 2010; 30:17035–40.
[13] Tam A, Dansereau C, Badhwar A, Orban P, Belleville S, Chertkow H, et al. Common effects of amnestic mild cognitive impairment on resting-state connectivity across four independent studies. Front Aging Neurosci 2015;7:242.
[14] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8:336–41.
[15] Bellec P, Benhajali Y, Carbonell F, Dansereau C, Albouy G, Pelland M, et al. Impact of the resolution of brain parcels on connectome-wide association studies in fMRI. Neuroimage 2015;123:212–28.
[16] Lancaster JL, Tordesillas-Gutiérrez D, Martinez M, Salinas F, Evans A, Zilles K, et al. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp 2007; 28:1194–205.
[17] Phipson B, Smyth GK. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. Stat Appl Genet Mol Biol 2010;9. Article39.
[18] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol 1995;57:289–300.
[19] Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. Activation likelihood estimation meta-analysis revisited. Neuroimage 2012; 59:2349–61.
[20] Eickhoff SB, Nichols TE, Laird AR, Hoffstaedter F, Amunts K, Fox PT, et al. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 2016;137:70–85.
[21] Li HJ, Hou XH, Liu HH, Yue CL, He Y, Zuo XN. Toward systems neuroscience in mild cognitive impairment and Alzheimer's disease: a meta-analysis of 75 fMRI studies. Hum Brain Mapp 2015; 36:1217–32.
[22] Jacobs HIL, Radua J, Lückmann HC, Sack AT. Meta-analysis of functional network alterations in Alzheimer's disease: toward a network biomarker. Neurosci Biobehav Rev 2013;37:753–65.
[23] Kim HJ, Cha J, Lee JM, Shin JS, Jung NY, Kim YJ, et al. Distinctive resting state network disruptions among Alzheimer's disease, subcortical vascular dementia, and mixed dementia patients. J Alzheimers Dis 2016;50:709–18.
[24] Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, Hedden T, et al. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. J Neurosci 2009;29:1860–73.
[25] Malek-Ahmadi M. Reversion from mild cognitive impairment to normal cognition: a meta-analysis. Alzheimer Dis Assoc Disord 2016;30:324–30.
[26] Köhler S, Hamel R, Sistermans N, Koene T, Pijnenburg YAL, van der Flier WM, et al. Progression to dementia in memory clinic patients without dementia: a latent profile analysis. Neurology 2013;81:1342–9.
[27] Jones DT, Knopman DS, Gunter JL, Graff-Radford J, Vemuri P, Boeve BF, et al. Cascading network failure across the Alzheimer's disease spectrum. Brain 2016;139:547–62.
[28] Adriaanse SM, Binnewijzend MA, Ossenkoppele R, Tijms BM, van der Flier WM, Koene T, et al. Widespread disruption of functional

brain organization in early-onset Alzheimer's disease. PLoS One 2014;9:e102995.
[29] Gour N, Felician O, Didic M, Koric L, Gueriot C, Chanoine V, et al. Functional connectivity changes differ in early and late-onset Alzheimer's disease. Hum Brain Mapp 2014;35:2978–94.
[30] Thomas JB, Brier MR, Bateman RJ, Snyder AZ, Benzinger TL, Xiong C, et al. Functional connectivity in autosomal dominant and late-onset Alzheimer disease. JAMA Neurol 2014;71:1111–22.
[31] Quiroz YT, Schultz AP, Chen K, Protas HD, Brickhouse M, Fleisher AS, et al. Brain imaging and blood biomarker abnormalities in children with autosomal dominant Alzheimer disease: a cross-sectional study. JAMA Neurol 2015;72:912–9.
[32] Chhatwal JP, Schultz AP, Johnson K, Benzinger TL, Jack C Jr, Ances BM, et al. Impaired default network functional connectivity in autosomal dominant Alzheimer disease. Neurology 2013;81:736–44.
[33] Sala-Llonch R, Fortea J, Bartrés-Faz D, Bosch B, Lladó A, Peña-Gómez C, et al. Evolving brain functional abnormalities in PSEN1 mutation carriers: a resting and visual encoding fMRI study. J Alzheimers Dis 2013;36:165–75.
[34] Shu H, Shi Y, Chen G, Wang Z, Liu D, Yue C, et al. Opposite neural trajectories of apolipoprotein E ε4 and ε2 alleles with aging associated with different risks of Alzheimer's disease. Cereb Cortex 2016;26:1421–9.
[35] Song H, Long H, Zuo X, Yu C, Liu B, Wang Z, et al. APOE effects on default mode network in Chinese cognitive normal elderly: relationship with clinical cognitive performance. PLoS One 2015;10:e0133179.
[36] Machulda MM, Jones DT, Vemuri P, McDade E, Avula R, Przybelski S, et al. Effect of APOE ε4 status on intrinsic network connectivity in cognitively normal elderly subjects. Arch Neurol 2011;68:1131–6.
[37] Goveas JS, Xie C, Chen G, Li W, Ward BD, Franczak MB, et al. Functional network endophenotypes unravel the effects of apolipoprotein E epsilon 4 in middle-aged adults. PLoS One 2013;8:e55902.
[38] Patel KT, Stevens MC, Pearlson GD, Winkler AM, Hawkins KA, Skudlarski P, et al. Default mode network activity and white matter integrity in healthy middle-aged ApoE4 carriers. Brain Imaging Behav 2013;7:60–7.
[39] Westlye ET, Lundervold A, Rootwelt H, Lundervold AJ, Westlye LT. Increased hippocampal default mode synchronization during rest in middle-aged and elderly APOE ε4 carriers: relationships with memory performance. J Neurosci 2011;31:7775–83.
[40] Su YY, Liang X, Schoepf UJ, Varga-Szemes A, West HC, Qi R, et al. APOE polymorphism affects brain default mode network in healthy young adults: a STROBE article. Medicine 2015;94:e1734.
[41] Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, Smith SM, et al. Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. Proc Natl Acad Sci U S A 2009;106:7209–14.
[42] Ferreira LK, Busatto GF. Resting-state functional connectivity in normal brain aging. Neurosci Biobehav Rev 2013;37:384–400.
[43] Evers EA, Klaassen EB, Rombouts SA, Backes WH, Jolles J. The effects of sustained cognitive task performance on subsequent resting state functional connectivity in healthy young and middle-aged male school teachers. Brain Connect 2012;2:102–12.
[44] Zuo XN, Kelly C, Di Martino A, Mennes M, Margulies DS, Bangaru S, et al. Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. J Neurosci 2010;30:15034–43.
[45] Scheinost D, Finn ES, Tokoglu F, Shen X, Papademetris X, Hampson M, et al. Sex differences in normal age trajectories of functional brain networks. Hum Brain Mapp 2015;36:1524–35.
[46] Podcasy JL, Epperson CN. Considering sex and gender in Alzheimer disease and other dementias. Dialogues Clin Neurosci 2016;18:437–46.
[47] Wang L, Brier MR, Snyder AZ, Thomas JB, Fagan AM, Xiong C, et al. Cerebrospinal fluid Aβ42, phosphorylated Tau181, and resting-state functional connectivity. JAMA Neurol 2013;70:1242–8.
[48] Elman JA, Madison CM, Baker SL, Vogel JW, Marks SM, Crowley S, et al. Effects of beta-amyloid on resting state functional connectivity within and between networks reflect known patterns of regional vulnerability. Cereb Cortex 2016;26:695–707.
[49] Wang Z, Zhang M, Han Y, Song H, Guo R, Li K. Differentially disrupted functional connectivity of the subregions of the amygdala in Alzheimer's disease. J Xray Sci Technol 2016;24:329–42.
[50] Sala-Llonch R, Bartrés-Faz D, Junqué C. Reorganization of brain networks in aging: a review of functional connectivity studies. Front Psychol 2015;6:663.
[51] He X, Qin W, Liu Y, Zhang X, Duan Y, Song J, et al. Abnormal salience network in normal aging and in amnestic mild cognitive impairment and Alzheimer's disease. Hum Brain Mapp 2014;35:3446–64.
[52] Zhou J, Greicius MD, Gennatas ED, Growdon ME, Jang JY, Rabinovici GD, et al. Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. Brain 2010;133:1352–67.
[53] Gour N, Ranjeva JP, Ceccaldi M, Confort-Gouny S, Barbeau E, Soulier E, et al. Basal functional connectivity within the anterior temporal network is associated with performance on declarative memory tasks. Neuroimage 2011;58:687–97.
[54] Heise V, Filippini N, Trachtenberg AJ, Suri S, Ebmeier KP, Mackay CE. Apolipoprotein E genotype, gender and age modulate connectivity of the hippocampus in healthy adults. Neuroimage 2014;98:23–30.
[55] Matura S, Prvulovic D, Butz M, Hartmann D, Sepanski B, Linnemann K, et al. Recognition memory is associated with altered resting-state functional connectivity in people at genetic risk for Alzheimer's disease. Eur J Neurosci 2014;40:3128–35.
[56] Trachtenberg AJ, Filippini N, Ebmeier KP, Smith SM, Karpe F, Mackay CE. The effects of APOE on the functional architecture of the resting brain. Neuroimage 2012;59:565–72.
[57] Mišić B, Betzel RF, Nematzadeh A, Goñi J, Griffa A, Hagmann P, et al. Cooperative and competitive spreading dynamics on the human connectome. Neuron 2015;86:1518–29.
[58] Sepulcre J, Schultz AP, Sabuncu M, Gomez-Isla T, Chhatwal J, Becker A, et al. In vivo tau, amyloid, and gray matter profiles in the aging brain. J Neurosci 2016;36:7364–74.
[59] Crossley NA, Mechelli A, Scott J, Carletti F, Fox PT, McGuire P, et al. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. Brain 2014;137:2382–95.
[60] Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001;54:1046–55.
[61] Jones DT, Vemuri P, Murphy MC, Gunter JL, Senjem ML, Machulda MM, et al. Non-stationarity in the "resting brain's" modular architecture. PLoS One 2012;7:e39731.
[62] Cheng W, Palaniyappan L, Li M, Kendrick KM, Zhang J, Luo Q, et al. Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. NPJ Schizophr 2015;1:15016.
[63] Wang K, Liang M, Wang L, Tian L, Zhang X, Li K, et al. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum Brain Mapp 2007;28:967–78.
[64] Zhang H-Y, Wang S-J, Xing J, Liu B, Ma Z-L, Yang M, et al. Detection of PCC functional connectivity characteristics in resting-state fMRI in mild Alzheimer's disease. Behav Brain Res 2009;197:103–8.
[65] Zhang H-Y, Wang S-J, Liu B, Ma Z-L, Yang M, Zhang Z-J, et al. Resting brain connectivity: changes during the progress of Alzheimer disease. Radiology 2010;256:598–606.
[66] Sheline YI, Raichle ME, Snyder AZ, Morris JC, Head D, Wang S, et al. Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. Biol Psychiatry 2010;67:584–7.

[67] Gili T, Cercignani M, Serra L, Perri R, Giove F, Maraviglia B, et al. Regional brain atrophy and functional disconnection across Alzheimer's disease evolution. J Neurol Neurosurg Psychiatry 2011;82:58–66.
[68] Wu X, Li R, Fleisher AS, Reiman EM, Guan X, Zhang Y, et al. Altered default mode network connectivity in Alzheimer's disease–a resting functional MRI and Bayesian network study. Hum Brain Mapp 2011;32:1868–81.
[69] Li R, Wu X, Fleisher AS, Reiman EM, Chen K, Yao L. Attention-related networks in Alzheimer's disease: a resting functional MRI study. Hum Brain Mapp 2012;33:1076–88.
[70] Damoiseaux JS, Prater KE, Miller BL, Greicius MD. Functional connectivity tracks clinical deterioration in Alzheimer's disease. Neurobiol Aging 2012;33:828.e19–30.
[71] Binnewijzend MAA, Schoonheim MM, Sanz-Arigita E, Wink AM, van der Flier WM, Tolboom N, et al. Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. Neurobiol Aging 2012;33:2018–28.
[72] Kenny ER, Blamire AM, Firbank MJ, O'Brien JT. Functional connectivity in cortical regions in dementia with Lewy bodies and Alzheimer's disease. Brain 2012;135:569–81.
[73] Zhu DC, Majumdar S, Korolev IO, Berger KL, Bozoki AC. Alzheimer's disease and amnestic mild cognitive impairment weaken connections within the default-mode network: a multi-modal imaging study. J Alzheimers Dis 2013;34:969–84.
[74] Balthazar MLF, Pereira FRS, Lopes TM, da Silva EL, Coan AC, Campos BM, et al. Neuropsychiatric symptoms in Alzheimer's disease are related to functional connectivity alterations in the salience network. Hum Brain Mapp 2014;35:1237–46.
[75] Yao H, Liu Y, Zhou B, Zhang Z, An N, Wang P, et al. Decreased functional connectivity of the amygdala in Alzheimer's disease revealed by resting-state fMRI. Eur J Radiol 2013;82:1531–8.
[76] Zhou B, Liu Y, Zhang Z, An N, Yao H, Wang P, et al. Impaired functional connectivity of the thalamus in Alzheimer's disease and mild cognitive impairment: a resting-state fMRI study. Curr Alzheimer Res 2013;10:754–66.
[77] Zhang Z, Liu Y, Zhou B, Zheng J, Yao H, An N, et al. Altered functional connectivity of the marginal division in Alzheimer's disease. Curr Alzheimer Res 2014;11:145–55.
[78] Weiler M, Fukuda A, Massabki LHP, Lopes TM, Franco AR, Damasceno BP, et al. Default mode, executive function, and language functional connectivity networks are compromised in mild Alzheimer's disease. Curr Alzheimer Res 2014;11:274–82.
[79] Balachandar R, John JP, Saini J, Kumar KJ, Joshi H, Sadanand S, et al. A study of structural and functional connectivity in early Alzheimer's disease using rest fMRI and diffusion tensor imaging. Int J Geriatr Psychiatry 2015;30:497–504.
[80] Pasquini L, Scherr M, Tahmasian M, Meng C, Myers NE, Ortner M, et al. Link between hippocampus' raised local and eased global intrinsic connectivity in AD. Alzheimers Dement 2015;11:475–84.
[81] Yi D, Choe YM, Byun MS, Sohn BK, Seo EH, Han J, et al. Differences in functional brain connectivity alterations associated with cerebral amyloid deposition in amnestic mild cognitive impairment. Front Aging Neurosci 2015;7:15.
[82] Sorg C, Riedl V, Mühlau M, Calhoun VD, Eichele T, Läer L, et al. Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. Proc Natl Acad Sci U S A 2007;104:18760–5.
[83] Bai F, Watson DR, Yu H, Shi Y, Yuan Y, Zhang Z. Abnormal resting-state functional connectivity of posterior cingulate cortex in amnestic type mild cognitive impairment. Brain Res 2009;1302:167–74.
[84] Bai F, Watson DR, Shi Y, Wang Y, Yue C, YuhuanTeng, et al. Specifically Progressive Deficits of Brain Functional Marker in Amnestic Type Mild Cognitive Impairment. PLoS One 2011;6:e24271.
[85] Xie C, Bai F, Yuan B, Yu H, Shi Y, Yuan Y, et al. Joint effects of gray matter atrophy and altered functional connectivity on cognitive deficits in amnestic mild cognitive impairment patients. Psychol Med 2015;45:1799–810.
[86] Jin M, Pelak VS, Cordes D. Aberrant default mode network in subjects with amnestic mild cognitive impairment using resting-state functional MRI. Magn Reson Imaging 2012;30:48–61.
[87] Han SD, Arfanakis K, Fleischman DA, Leurgans SE, Tuminello ER, Edmonds EC, et al. Functional connectivity variations in mild cognitive impairment: associations with cognitive function. J Int Neuropsychol Soc 2012;18:39–48.
[88] Liang P, Wang Z, Yang Y, Li K. Three subsystems of the inferior parietal cortex are differently affected in mild cognitive impairment. J Alzheimers Dis 2012;30:475–87.
[89] Hahn K, Myers N, Prigarin S, Rodenacker K, Kurz A, Förstl H, et al. Selectively and progressively disrupted structural connectivity of functional brain networks in Alzheimer's disease - revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence. Neuroimage 2013;81:96–109.
[90] Myers N, Pasquini L, Göttler J, Grimmer T, Koch K, Ortner M, et al. Within-patient correspondence of amyloid-β and intrinsic network connectivity in Alzheimer's disease. Brain 2014;137:2052–64.
[91] Koch K, Myers NE, Göttler J, Pasquini L, Grimmer T, Förster S, et al. Disrupted Intrinsic Networks Link Amyloid-β Pathology and Impaired Cognition in Prodromal Alzheimer's Disease. Cereb Cortex 2015;25:4678–88.
[92] Das SR, Pluta J, Mancuso L, Kliot D, Yushkevich PA, Wolk DA. Anterior and posterior MTL networks in aging and MCI. Neurobiol Aging 2015;36 Suppl 1:S141–50.e1.
[93] Gardini S, Venneri A, Sambataro F, Cuetos F, Fasano F, Marchi M, et al. Increased functional connectivity in the default mode network in mild cognitive impairment: a maladaptive compensatory mechanism associated with poor semantic memory performance. J Alzheimers Dis 2015;45:457–70.

Appendix B

Subtypes of functional brain connectivity as early markers of neurodegeneration in Alzheimer's disease

Pierre Orban, Angela Tam, Sebastian Urchs, Mélissa Savard, Cécile Madjar, AmanPreet Badhwar, Christian Dansereau, Jacob Vogel, Amir Shmuel, Alain Dagher, Sylvia Villeneuve, Judes Poirier, Pedro Rosa-Neto, John Breitner, Pierre Bellec

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Subtypes of functional brain connectivity as early markers of neurodegeneration in Alzheimer's disease

Pierre Orban[1,2,3,4], Angela Tam[1,2,6], Sebastian Urchs[1,6], Melissa Savard[2], Cécile Madjar[2], AmanPreet Badhwar[1], Christian Dansereau[1,5], Jacob Vogel[2,6], Amir Schmuel[6], Alain Dagher[6], Sylvia Villeneuve[2,6], Judes Poirier[2,6], Pedro Rosa-Neto[2,6], John Breitner[2,6], Pierre Bellec[1,5], for the Alzheimer's Disease Neuroimaging Initiative* and the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease Program**

1. Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montreal, QC, Canada
2. Douglas Mental Health University Institute, Research Centre, Montreal, QC, Canada
3. Centre de recherche de l'Intitut universitaire de santé mentale de Montréal, Montreal, QC, Canada
4. Département de psychiatrie, Université de Montréal, Montreal, QC, Canada
5. Département d'informatique et recherche opérationnelle, Université de Montréal, Montreal, QC, Canada
6. McGill University, Montreal, QC, Canada

* Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

** Data used in preparation of this article were obtained from the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) program (http://www.prevent-alzheimer.ca), data release 2.0 (November 30, 2015). As such, the investigators of the PREVENT-AD program contributed to the design and implementation of PREVENT-AD and/or provided data but did not participate in analysis or writing of this report. A complete listing of PREVENT-AD Research Group can be found in: https://preventad.loris.ca/acknowledgements/acknowledgements.php?date=[2015-11-30].

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Highlights

- Reliable functional brain network subtypes accompany cognitive impairment in AD
- Symptom-related subtypes exist in the default-mode, limbic and salience networks
- A limbic subtype is associated with a familial risk of AD in healthy older adults
- Limbic subtypes also associate with beta amyloid deposition and ApoE4

In Brief

We found reliable subtypes of functional brain connectivity networks in older adults, associated with AD-related clinical symptoms in patients as well as several AD risk factors/biomarkers in asymptomatic individuals.

Summary

The heterogeneity of brain degeneration has not been investigated yet for functional brain network connectivity, a promising biomarker of Alzheimer's disease. We coupled cluster analysis with resting-state functional magnetic resonance imaging to discover connectivity subtypes in healthy older adults and patients with cognitive disorders related to Alzheimer's disease, noting associations between subtypes and cognitive symptoms in the default-mode, limbic and salience networks. In an independent asymptomatic cohort with a family history of Alzheimer's dementia, the connectivity subtypes had good test-retest reliability across all tested networks. We found that a limbic subtype was overrepresented in these individuals, which was previously associated with symptoms. Other limbic subtypes showed associations with cerebrospinal fluid $A\beta_{1-42}$ levels and ApoE4 genotype. Our results demonstrate the existence of reliable subtypes of functional brain networks in older adults and support future investigations in limbic connectivity subtypes as early biomarkers of Alzheimer's degeneration.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Introduction

Alzheimer's disease (AD) is a chronic neurodegenerative condition that gives rise to the most common form of dementia, with severe memory and cognitive impairments. Importantly, the clinical expression of AD becomes apparent only decades after the development of neuropathological processes, such as the accumulation of amyloid-β (Aβ) plaques and tau neurofibrillary tangles. The long preclinical buildup of AD pathology presents an opportunity to prevent, rather than repair, neurodegeneration (Dubois et al., 2016; Sperling et al., 2012). Functional brain connectivity measured with resting-state functional magnetic resonance imaging (rs-fMRI) may capture early synaptic dysfunction in AD (Selkoe, 2002; Tampellini, 2015) and is thus a promising candidate biomarker for AD (Badhwar et al., 2017; Brier et al., 2014; Jones et al., 2016; Vemuri et al., 2012). However, the current literature has largely relied on comparisons between group averages of patients and cognitively healthy individuals. Such cross-sectional analyses neglect the considerable phenotypic heterogeneity present both in patient and control populations. The primary objective of this work was to characterize the heterogeneity of functional brain connectivity in older adults, and identify network subtypes associated with AD at the clinical and preclinical stages.

A prevalent model of AD postulates that symptoms arise as a consequence of disruptions in distributed networks, rather than local, circumscribed alteration in neural processing (Delbeuck et al., 2003; Seeley et al., 2009). The seminal work of (Greicius et al., 2004) in symptomatic AD demonstrated alterations in functional brain connectivity in the so-called default-mode network (DMN), whose topography overlaps substantially with patterns of end-stage Aβ deposition (Buckner et al., 2005). A recent meta-analysis concluded to convergent evidence across over 30 publications looking at functional brain connectivity in clinical cohorts, i.e. patients with mild cognitive impairment or AD dementia, and confirmed the DMN as a key affected brain component (Badhwar et al., 2017). Connectivity disturbances in other large-scale brain networks were also found consistently, in particular in the limbic and salience networks. At a preclinical stage, rs-fMRI connectivity has been shown to be impacted in cognitively healthy older adults at risk of AD due to abnormal levels of cerebrospinal fluid (CSF) $A\beta_{1-42}$ or tau proteins (Jiang et al., 2016; Wang et al., 2013), increased cerebral Aβ deposits (Elman et al., 2016), and presence of apolipoprotein E ε4 allele - ApoE4 (Sheline et al., 2010), the major genetic risk factor in sporadic AD. A familial history of sporadic AD in first-degree relatives is the second most important risk factor of AD (Tanzi, 2012), and was shown to impact DMN connectivity even in ApoE4 non-carriers, thus highlighting additional genetic risk factors (Wang et al., 2012).

Despite mounting evidence of rs-fMRI as an early marker of AD, the current literature neglects the considerable heterogeneity present in both patients and controls. Post-mortem histological examination of AD pathology in brain tissue samples (Hyman et al., 2012) indeed does not align closely with clinical diagnosis. Between 30 and 50% of patients diagnosed with AD dementia in fact do not present Alzheimer's pathology, depending on the level of neuropathological confidence (Beach et al., 2012).

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Conversely, the same study reported that close to 40% of patients diagnosed with non-AD dementia show minimal signs of AD pathology. Some cognitively healthy persons included in control groups may also suffer from preclinical AD, with 10% to 30% of them having Aβ deposition in their brain (Chételat et al., 2013), and some of them exhibiting high loads of neurofibrillary tangles (Mufson et al., 2016). Data-driven analysis of structural MRI subtypes in AD further showed that symptomatic heterogeneity (Belleville et al. 2007; Scheltens et al. 2016) is at least partly related to different modes of atrophy spreading in AD (Dong et al., 2017; Zhang et al., 2016). Complementary to subtypes of brain atrophy, a recent work (Doan et al., 2017) showed that connectivity subtypes can also be observed using diffusion magnetic resonance imaging, in patients suffering from AD dementia, MCI or subjective cognitive impairment, with subtypes accompanying the severity of cognitive impairment. The established heterogeneity in structural brain degeneration calls to re-examine the current evidence for rs-fMRI as an AD biomarker using a subtyping approach.

The overarching goal of the present work was to identify one or multiple subtypes of functional brain connectivity associated with AD, either at a clinical or preclinical stage. We first applied a data-driven cluster analysis to identify subgroups of subjects with homogeneous subtypes of brain connectivity within a mixed cohort of 130 subjects, including patients with AD dementia (AD subjects, N = 21), patients with mild cognitive impairment (MCI subjects, N= 44), and elderly healthy controls (HC subjects, N= 65) (Figure 1, Table 1). This mixed cohort, referred to as the ADNI2-MTL sample,poolsdatafrom 2 sites of ADNI2 (ADNI2a and ADNI2b) and 3 studies conducted at Montreal sites (MNI, CRIUGMa and CRIUGMb), in an attempt to extract robust subtypes that will generalize well to new studies (Orban et al., 2017). For each brain network and connectivity subtype, we tested whether a particular subtype was associated with the presence of mild or severe symptoms. We then investigated if the

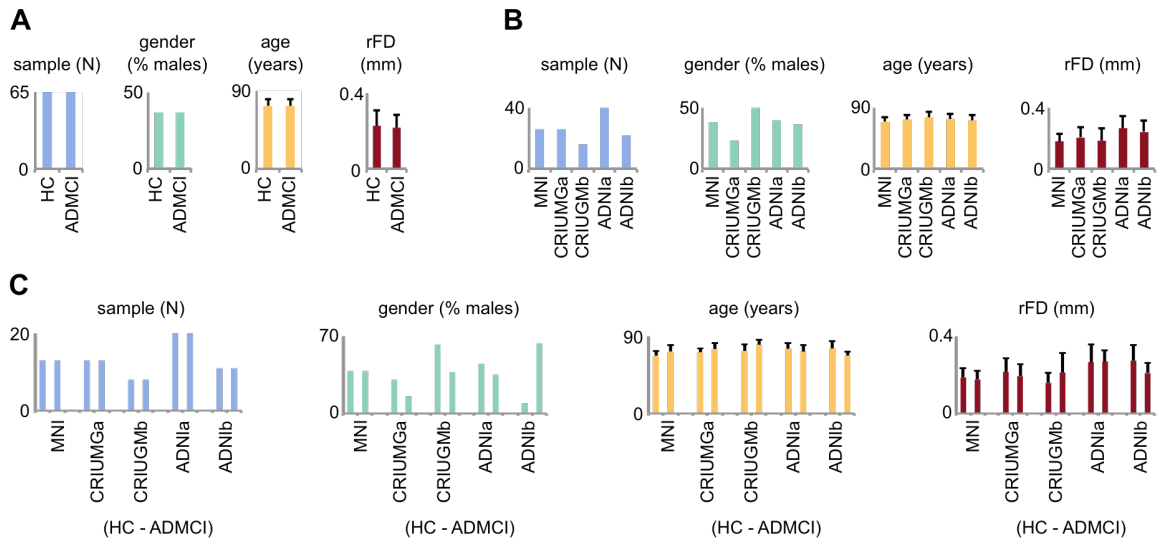

Figure 1. Matching between ADMCI patients and HC
(A) Patients and controls were matched with respect to sample size, gender, age and motion levels after scrubbing (residual frame displacement, rFD). (B) Between-site differences on4 such variables are shown irrespective of clinical status. (C) The number of patients and controls are perfectly balanced within sites.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

subtype membership was a reliable quantity using test-retest data in an independent sample of 231 cognitively healthy older adults, with a familial history of AD (FH subjects) (Orban et al., 2015). As those subjects are at risk for AD, we tested if the subtypes associated with symptoms would already be overrepresented in the asymptomatic FH cohort. We further tested in these FH subjects the association between functional subtypes and known biomarkers/risk factors of AD, namely CSF $A\beta_{1-42}$, Tau levels as well as ApoE4 genotype.

Results

Subtypes of functional brain networks

To identify subtypes of functional brain networks, we first generated individual functional connectivity maps for seven large-scale networks together covering the entire brain (Figure 2A-B). These reference networks were obtained from an independent dataset from 200 healthy young subjects (Bellec et al. 2015), and were labeled as cerebellar, limbic, motor, visual, default-mode, fronto-parietal and salience networks. For each network, a hierarchical cluster analysis was applied on 130 individual network maps from the ADNI2-MTL dataset, after regression of phenotypic and site confounds, in order to identify subgroups of subjects with homogeneous brain maps. Visual inspection suggested the presence of at least three voxelwise connectivity subgroups (Figure 2C-D). A brain map averaged across all subjects within a subgroup defined a subtype of network connectivity, highlighting specific brain areas that differed between that subgroup and the overall population average (Figure 2E). Subtype maps revealed high connectivity with their reference network, yet also exhibited noticeable variations. These differences were not only observed in the associated network (within-network connectivity) but also in other brain areas (between-network connectivity). For instance, subtypes of the DMN could be distinguished from one another not only in terms of connectivity levels within the precuneus or anterior medial prefrontal cortex, two key nodes of the default-mode network, but also with regards to connectivity strength in the anterior cingulate, associated with the salience network. For each network, we generated the spatial correlations between individual connectivity maps and each average subtype map, hereafter referred to as weights (Figure 2F). These continuous subtype weights revealed that some individual maps were highly correlated with the subtypes, while others had only milder correlations, sometimes of similar amplitude for different subtypes. The subtype decomposition was therefore a discrete approximation of a continuous distribution of individual maps, rather than a set of clear-cut entities.

A comparison of clustering outcomes for the seven networks revealed that 3 subgroups of subjects at least could be evidenced in all networks (Figure 3A). As observed for the DMN, subtype maps showed distributed variations inside and outside the network of reference for all networks. While between-subject correlation values had similar amplitudes across networks, the size of the subgroups varied from one network to another. We tested the correspondence of subject clustering solutions between

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

networks by computing the adjusted rand index (ARI) for all pairwise comparisons (Figure 3B). The near-chance level of this metric (0.04 ± 0.04) demonstrated that subjects with similar connectivity maps for a given network did not have particularly similar maps for other networks, thus highlighting heterogeneity in functional brain connectivity patterns.

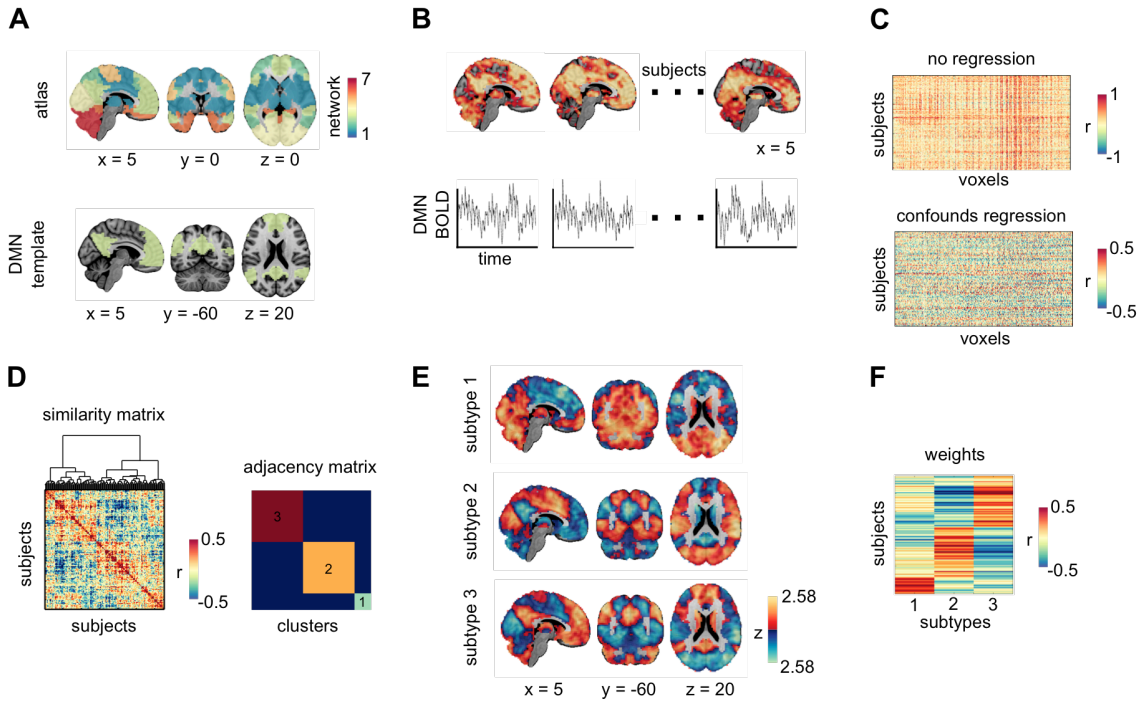

Figure 2. Extraction of subtypes and weights
(A) Functional subtypes were identified separately for 7 networks delineated at the whole-brain level in an independent sample of healthy subjects. The procedure is shown for the default-mode network (DMN). (B) Network-based connectivity maps were computed for each subject through the correlation of every voxel's time course of activity with the average signal in the reference network. (C) Site, gender, age and motion were regressed out from functional connectivity maps across subjects. (D) A hierarchical cluster analysis was conducted to identify 3 homogeneous subgroups of subjects with similar connectivity maps. (E) Difference subtypes show how the average connectivity maps of each separate subgroup of subjects differ from the grand average. (F) Weights consisted in correlations between the connectivity maps of every subject with that of each subtype.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

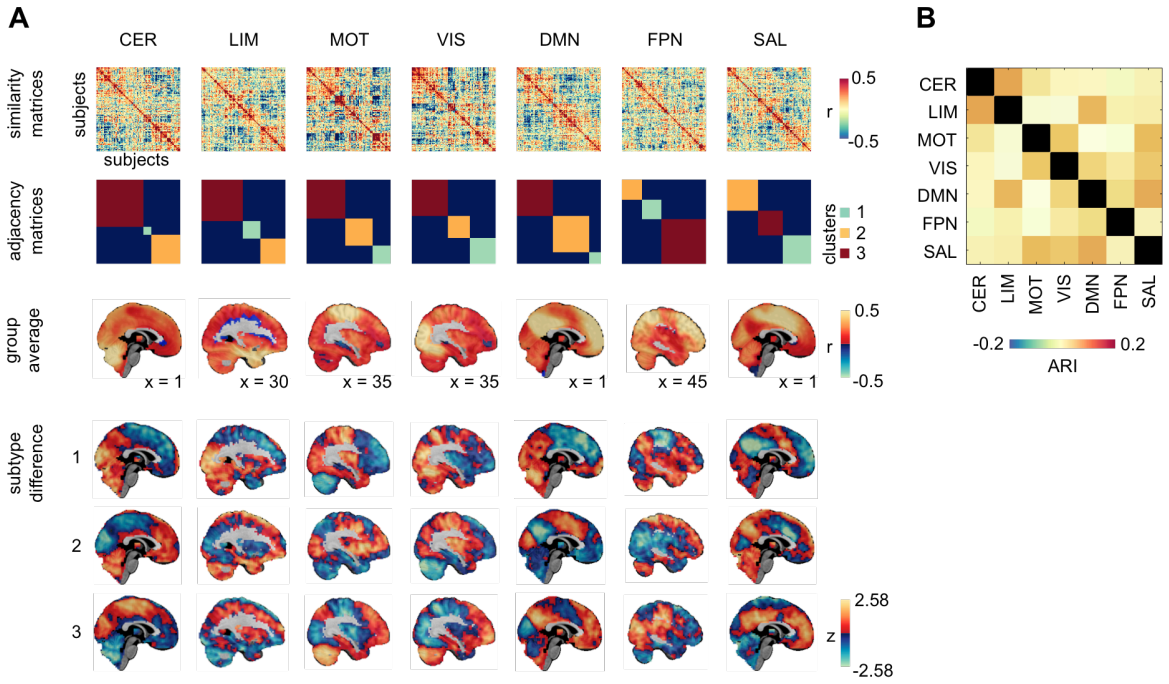

Figure 3. Correspondence of cluster (subtype) solutions across networks
(A) For each of the 7 networks (columns) are given the similarity matrix that shows the similarity of network connectivity maps between all pairs of subjects (first row), the adjacency matrix that reveals homogeneous subgroups of subjects identified by cluster analysis (second row), the average network connectivity map for all subjects (third row), and the difference subtype connectivity maps obtained by differences between the group average and the average connectivity maps for each subgroup of subjects (fourth to sixth rows). (B) The adjusted rand index (ARI) reveals the correspondence of subject clustering solutions between all pairs of networks. CER, cerebellum; LIM, limbic; MOT, motor; VIS, visual; DMN, default-mode; FPN, fronto-parietal; SAL, salience.

Brain network subtypes are associated with clinical symptoms

Given the observation that subtypes reflected both continuous and discrete phenomena, we adopted a dual statistical evaluation of their association with clinical symptoms in ADMCI subjects (Figure 4). In the former case, differences in average subtype weights between ADMCI and HC were assessed independently for each subtype of the seven reference networks, using a linear regression model. Significant associations were found for one limbic, two default-mode and two salience subtypes ($q < 0.05$ with FDR correction over 21 network subtypes), in line with our expectations. An uncorrected effect was also seen for an additional limbic subtype ($p < 0.05$). Effects were of medium size ($0.09 <$ Cohen's $f2 < 0.25$). Of these six subtypes, half of the associations with symptoms were positive (i.e. higher average weight load in ADMCI persons) and the remainder negative (i.e. lower average weight load in ADMCI patients). Instances of positive and negative associations with symptoms were observed in all three aforementioned networks.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

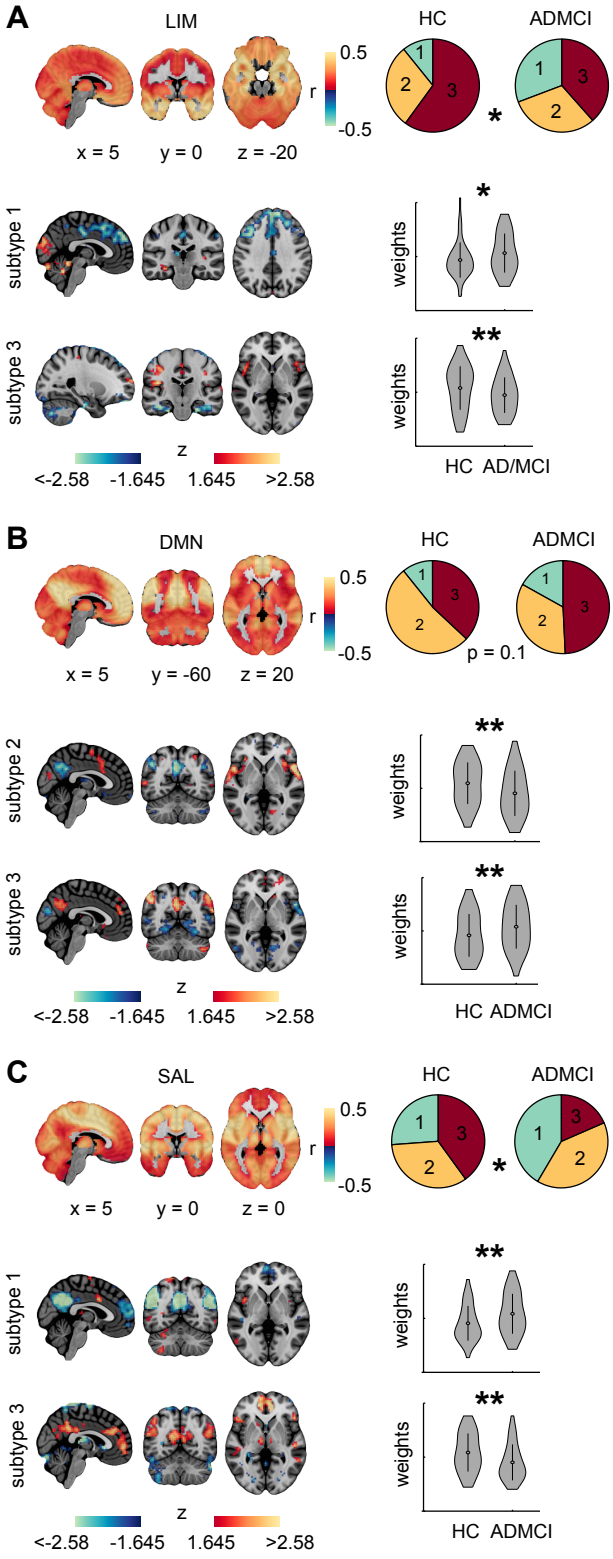

Figure 4. Functional network subtypes associated with clinical symptoms
Significant associations with ADMCI were found in the limbic (A), default-mode (B) and salience (C) networks. For each network are shown the group average connectivity map and the connectivity subtypes that are significantly more or less present in ADMCI patients than controls (difference maps are given). Pie charts report the distributions of subjects across subtypes in each group. Violin plots show the distribution of weights in the two groups for each subtype with a significant association. ** and * respectively denote significance at $q^{FDR}<0.05$ and $p<0.05$ (uncorrected).

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

A general observation was that subtypes positively associated with symptoms (PAS) had increased within-network connectivity but decreased between-network connectivity as compared to sample averages of networks. The PAS limbic subtype was notably defined by increased hippocampal connectivity (within-network) but decreased connectivity in dorsomedial prefrontal areas located in the DMN (between-network). An inverse pattern was seen in subtypes negatively associated with symptoms (NAS). The NAS limbic subtype had decreased connectivity in the hippocampus but increased connectivity in the insula. Subtypes of the default-mode and salience network provided mirror pictures of PAS and NAS connectivity profiles. Decreased connectivity in the posterior cingulate and medial prefrontal region relative to the sample average was NAS for the default-mode network but PAS for the salience network. Similarly, decreased connectivity in the insula and anterior cingulate cortex compared to the sample average was evidenced to be NAS for the salience network but PAS for the default-mode network.

Statistics on discrete effects provided concordant effects at uncorrected thresholds. For each network, we evaluated with Chi2 tests whether ADMCI and HC subjects were distributed unevenly across subtypes. Unequal distributions were seen for the limbic ($p < 0.05$), default-mode ($p = 0.1$) and salience ($p < 0.05$) networks. Effect sizes were in the small-to-moderate range, with Cramer's V values of 0.27, 0.19 and 0.24 in the limbic, default-mode and salience networks, respectively.

Connectivity maps in FH subjects are reproducibly matched to subtypes from the clinical cohort

We assessed the reliability of matching connectivity maps in FH subjects from the PREVENT-AD cohort with the subtypes defined in the MTL-ADNI2. We thus generated individual functional connectivity maps separately for two runs, in each of the seven networks. Weights were computed for individual network maps, indicating their similarity with each of the 21 network subtypes previously defined in the MTL-ADNI2 sample (Figure 5). Intraclass correlations (ICC) indicated a fair-to-good correspondence of subtype weights between runs. Weights of all network subtypes had ICC values > 0.45 (max = 0.68, mean = 0.56), except for the PAS salience subtype (0.29). The default-mode and limbic PAS subtype weights had ICCs of 0.50 and 0.55, respectively.

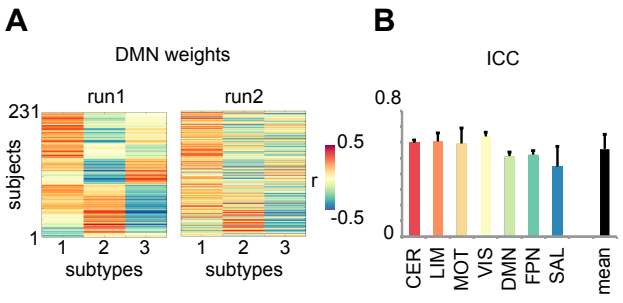

Figure 5. Reliability of subtype matching in FH subjects
(A) Matching of connectivity maps in FH subjects with subtypes found in the mixed population of ADMCI patients and controls is shown for the DMN in two separate runs. (B) Test-retest between runs was determined with intra-class correlation (ICC), showing fair-to-good correspondence across networks and subtypes.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Subtypes are associated with biomarkers of AD in FH subjects

We next examined the possibility that cognitively healthy FH older adults already exhibited PAS subtypes, and more so than typical healthy elderly individuals. Individual functional connectivity maps were averaged for the two separate runs in 231 FH subjects from the PREVENT-AD cohort. For each of the three networks found to be associated with clinical symptoms, FH subjects were matched to network subtypes defined in the MTL-ADNI sample based on maximal weights. Distributions of FH subjects across subtypes were not significantly different than those of either ADMCI or HC participants in the default-mode and salience networks (Figure 6A). However, proportions of FH subjects across limbic subtypes differed significantly from those of typical HC older adults ($q < 0.05$) but not from ADMCI patients ($p = 0.9$).

The idea that connectivity subtypes might reflect a covert pathological AD process in cognitively healthy elderly individuals would be reinforced by the observation that such connectivity profiles correlate with known biomarkers of AD. We thus further investigated the relationship between connectivity subtypes and APOE genotype ($N = 228$) as well as CSF levels of $A\beta_{1\text{-}42}$, tTau and pTau ($N = 59$) (Figure 6C). Surprisingly, APOE allele 4 carriers showed less association than non carriers with the limbic PAS subtype ($q < 0.05$), with a small effect size (Cohen's $f2 = 0.04$). However, findings consistent with predictions were observed for CSF Aß42 levels and another limbic subtype. Subjects with high levels of CSF $A\beta_{1\text{-}42}$ had limbic connectivity maps that resembled more the NAS limbic network ($q < 0.05$; Cohen's $f2 = 0.13$). Low levels of CSF $A\beta_{1\text{-}42}$ were associated with another limbic subtype that shared some similarities with the PAS limbic subtype, for instance increased hippocampal connectivity ($q < 0.05$; Cohen's $f2 = 0.1$). No associations were found between Tau or pTau CSF levels and any subtype of either the limbic, default-mode or salience networks.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

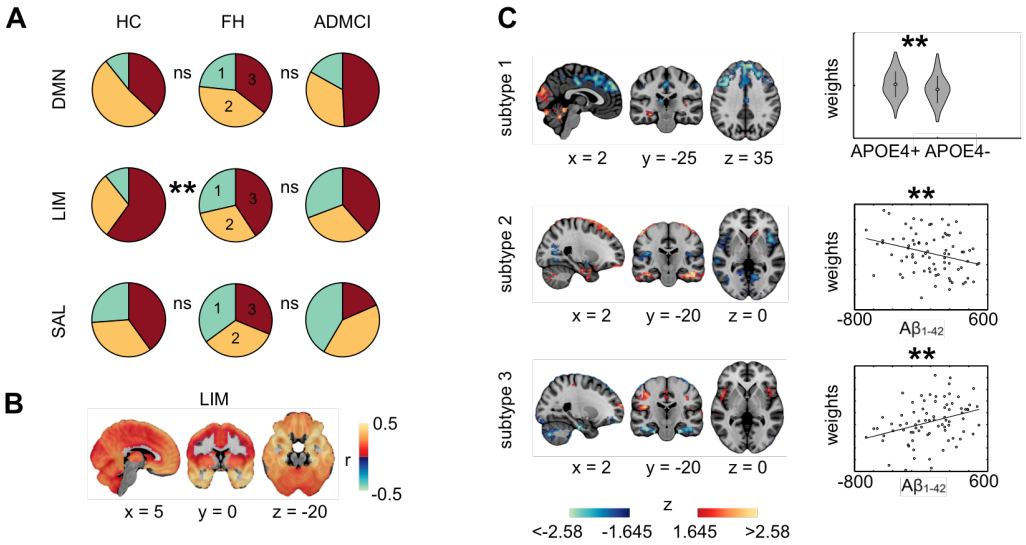

Figure 6. Connectivity subtypes in FH subjects
(A) Pie charts show that FH subjects differ from controls but not ADMCI patients in their distribution across subtypes for the limbic network (B). (C) Three distinct limbic network subtypes show either positive or negative associations with ApoE4 status or CSF $A\beta_{1-42}$ levels. ** denotes significance at $q^{FDR}<0.05$.

Discussion

Capturing heterogeneity through subtyping

Our subtyping approach was motivated by the lack of specificity and sensitivity of a clinical diagnosis of AD dementia against a histopathological diagnosis of AD pathology (Beach et al., 2012) and the variability of cognitive and neurobiological alterations in AD (Lam et al., 2013; Scheltens et al., 2016). As done previously for structural atrophy patterns (Dong et al., 2017; Hwang et al., 2016; Zhang et al., 2016) and white matter structural dysconnectivity (Doan et al., 2017), we employed a subtype analysis that identified subgroups of subjects sharing similar functional brain connectivity, in a fully data-driven way and irrespective of clinical diagnosis. This is an important conceptual difference with more traditional cross-sectional comparisons between clinical cohorts, which assumes some homogeneity in connectivity within each group, e.g. (Badhwar et al., 2017, 2016; Jones et al., 2016; Korolev et al., 2016). Improved characterization of the inherent heterogeneity of brain dysconnectivity in AD will ultimately facilitate more personalized diagnosis and treatment. This new line of inquiry is made possible by large neuroimaging databases such as the ADNI, and will become increasingly important with the emergence of populational cohorts with associated neuroimaging repositories, such as the UK biobank (Miller et al., 2016).

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Association between connectivity subtypes and clinical symptoms

Using rs-fMRI, we identified functional brain connectivity subtypes associated both positively and negatively with symptoms. A variety of causal mechanisms may explain such associations, which may co-exist. An association may reflect the direct progression of AD neurodegeneration in the brain (Jones et al., 2016), the presence of comorbidities (Profenno et al., 2010), as well as some form of cognitive reserve, or lack thereof (Stern, 2006). The existence of an association in itself is not enough to disambiguate between these different interpretations. Associations between connectivity subtypes and symptoms were selectively detected in the default-mode, salience and limbic networks. These three networks have consistently been reported in the literature as altered in patients with AD dementia or MCI, see (Badhwar et al., 2017; Vemuri et al., 2012) for reviews. The associated subtype maps pointed at changes both within networks, e.g. higher intra-network connectivity in PAS DMN subtype, and between networks, e.g. decreased inter-network connectivity in PAS DMN subtypes with regions of the salience network.

Translation of connectivity subtypes from clinical to non-clinical individuals

The distribution of connectivity subtypes in a group of cognitively normal FH individuals was found to resemble more the subtype distribution of a patient group than that of a control group. This observation was made only for the limbic network, but not the default-mode and salience networks. Assuming functional connectivity subtypes partly reflect the progression of AD pathology, finding early dysconnectivity in the limbic network is consistent with the Braak staging of neurodegeneration (Braak and Braak, 1991) and the increased risk of sporadic AD due to family history (Tanzi, 2012). Conversely, the limbic subtype negatively associated with symptoms was under-represented in FH individuals, and was shown to positively associate with CSF $A\beta_{1\text{-}42}$ levels. Taken together, these associations support the notion that different subtypes of limbic connectivity reflect the progression of AD pathophysiology at a preclinical stage. A finding that was more difficult to interpret was that ApoE4 carriers had significantly less weight on the limbic subtype positively associated with symptoms. With previous literature on ApoE4 and resting-state connectivity sometimes reporting contradictory findings (Filippini et al., 2009; Sheline et al., 2010), we believe longitudinal data on a large cohort would be necessary to clarify the relationships among resting-state connectivity, Aβ deposition and ApoE4 status.

Generalization of brain connectivity subtypes across datasets

The translation of connectivity across cohorts raises the question of generalization across scanning sites. Research has indeed indicated that multisite scanning generates substantial site-specific bias in connectivity measures (Dansereau et al., 2017; Yan et al., 2013). In our multisite clinical sample, we took great care to control for confounding site effects on brain connectivity subtypes. The identification of network subtype was thus invariant to scanning site to a large extent. However, the cohort of individuals at

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

risk of AD due to their familial history was entirely scanned at a single site. The fact that we found associations with known biomarkers or risk factors of AD specifically in the limbic network supports that brain connectivity subtypes are fairly robust to site effects. Subtype weights also had good test-retest reliability in the PREVENT-AD cohort, although the subtype maps were generated on ADNI-MTL. Important areas for future work will be to identify imaging protocols that further minimize differences in brain connectivity subtypes across scanners.

Finer subtypes

Groups of patients that defined subtypes did not overlap a lot across networks, including for subtypes positively associated with symptoms. There is thus some degree of independence between subtypes across networks, possibly reflecting heterogeneity of neurodegeneration across patients. Even though we estimated only 3 subtypes per network, there are still a very large number of possible combinations of subtypes across 7 networks. Subtype maps being an average of a subgroup of subjects, a minimum number of 20 subjects seems warranted to stabilize the subtype maps. The total sample size of our discovery dataset thus constrained the maximal number of subtypes we could feasibly investigate. We thus decided to use low numbers of subtypes and networks for this first evaluation of the feasibility of functional subtypes in AD, yet higher numbers could be explored in a larger sample.

Multi-network and multimodal subtypes

A natural extension of this work would be to integrate subtypes across multiple networks, imaging modalities and measures into a single predictor of AD status. Associations with clinical symptoms or AD biomarkers reported here had weak to moderate effect sizes, despite reaching statistical significance. Recent state-of-the-art model of progression from MCI to dementia indeed merge biomarkers across multiple domains, including cognitive evaluations, imaging and plasma markers (Korolev et al., 2016). High-dimensional imaging biomarkers such as structural and diffusion MRI are amenable to subtyping (Doan et al., 2017; Hwang et al., 2016; Zhang et al., 2016). We believe that subtyping could be used in the near future to identify a highly accurate multimodal predictor of AD, both for diagnosis and prognosis purposes. Resting-fMRI will likely contribute to such a multimodal predictor, as it is uniquely sensitive to brain function, at least compared to other MRI modalities. Our findings suggest that limbic subtypes in particular are promising biomarkers for the purpose of early AD diagnosis.

Conclusions

The present work demonstrates that rs-fMRI can be used to subtype the heterogeneity of functional networks in older adults. We found that subtypes have a good test-retest reliability and associate with symptoms in patients suffering from MCI or AD dementia. We also found that subtypes associate with various biomarkers and risk factors of AD in cognitively normal individuals: familial history of AD dementia, beta amyloid deposition,

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

ApoE4 status. Our findings support the notion that rs-fMRI subtypes are sensitive to AD progression up to the preclinical stage, and may contribute to future efforts towards an accurate early diagnosis of AD using multimodal biomarkers.

Experimental procedures

Participants

The MTL-ADNI2 multisite sample aggregated data from 5 different studies: 3 samples from the Montreal area (one from the Montreal Neurological Institute, MNI, and two from the *Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal*, CRIUGMa and CRIUMGb), and 2 samples with distinct acquisition protocols from the Alzheimer's Disease Neuroimaging Initiative 2 (ADNI2a and ADNI2b) (Table 2). We selected subsamples of the MNI, CRIUGMa, CRIUGMb, ADNI2a and ADNI2b datasets such that patients and controls groups had identical sample size for each acquisition protocol or study, respectively 13, 13, 8, 20 and 11 subjects per group. The combined sample included 65 patients diagnosed with either amnestic MCI or AD dementia and 65 cognitively normal controls. Patients and controls were selected from a larger initial pool such that they would be matched for age, gender ratio as well as motion (see rs-fMRI preprocessing section). Distributions of age, gender and motion were as follows for patients vs. controls: age (mean ± std) = 72.7 ± 7.9 vs. 72.6 ± 7.3 years old, 41/24 vs. 41/24 females/males, residual frame displacement (mean ± std) = 0.22 ± 0.07 vs. 0.23 ± 0.08. All subjects gave informed consent to participate in these studies, which were approved by the research ethics committees of the institutions involved in data acquisition. Consent was obtained for data sharing and secondary analysis, the latter being approved by the ethics committee at the CRIUGM.

The PREVENT-AD dataset used in the present analysis included 231 cognitively healthy older adults with a known family history of AD, as reflected by a diagnosis of AD dementia in parent or first-degree relatives. PREVENT-AD participants were younger (mean ± std: 64.1 ± 5.7 years old) than subjects in the MTL-ADNI2 multisite sample and were not balanced for gender (172/59 females/males). All subjects had given informed consent and the study was approved by the "Research, Ethics and Compliance Committee" of McGill University.

Note on the cohorts

The ADNI2 data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership representing efforts of many co-investigators from a broad range of academic institutions and private corporations. A

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

central goal of ADNI is to facilitate the discovery of biomarkers of very early AD progression, using MRI among other techniques. ADNI was followed by ADNI-GO and ADNI-2. In this study, we only included subjects from the two ADNI2 scanners (Achieva and Intera) associated with the largest samples. For up-to-date information, see www.adni-info.org.

The PREVENT-AD data were obtained from the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) program data release 2.0 (November 30, 2015). The cohort of this program was composed of cognitively healthy individuals at increased risk of AD dementia because they have / had a first-degree relative (parent or sibling) who has / had dementia suggestive of AD. This cohort includes volunteers of age 60 or older (55 or older if current age is within 15 years of affected relative's estimated age at onset of dementia). One current project consists of an observational study where participants are followed longitudinally once a year with a battery of tests and imaging modalities. In the present work, we focused on baseline data. A subset of test-retest rsfMRI data in 80 PREVENT-AD subjects has been shared publicly (Orban et al., 2015).

Clinical evaluation

All subjects from the MTL-ADNI2 and PREVENT-AD samples underwent neuropsychological testing to assess cognitive function, including memory, language and executive abilities. However, the neuropsychological tests administered to participants varied across sites, as did criteria and clinical scales used for diagnosis of either MCI or AD. Briefly, patients with (amnestic) MCI had memory complaints and objective cognitive loss, yet showed intact functional abilities and did not meet criteria for a diagnosis of dementia in contrast with AD patients. HC demonstrated intact cognitive functions. Details on clinical evaluation for each cohort per site follow.

In ADNI2, the Mini-Mental State Evaluation (MMSE) and Clinical Dementia Rating (CDR) were used to distinguish between HC, MCI and AD subjects. MMSE scores were inclusively comprised between 24-30, 24-30 and 20-26 for HC, MCI and AD subjects, respectively. MCI patients had a CDR of 0.5 and AD patients a CDR of 0.5 or 1. An objective memory loss was evidenced with the Wechsler Memory Scale Logical Memory II in MCI, yet other cognitive domains and functional activities were unaffected. In addition, there was an absence of dementia, by contrast with AD patients who met the National Institute of Neurological and Communicative Disorders and Stroke / Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD (McKhann et al. 1984). The MNI sample only included MCI patients, who were similarly diagnosed using the MMSE, following Petersen Criteria (Petersen 2004). Subjects in the CRIUGM samples were administered the MMSE as well as the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) and the Mattis Dementia Rating Scale (Schmidt et al. 1994). The diagnosis of MCI was made based on scores equal to or >1.5 standard deviations below the mean adjusted for age and education on memory tests, with input from a neurologist. A diagnosis of AD was determined according to the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; American Psychiatric Association, 2000) and NINCDS/ADRDA clinical criteria, with input from a neurologist. Participants in the PREVENT-AD were evaluated for any

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

cognitive impairment and symptoms suggestive of AD using the Repeatable Battery for the Assessment of Neuropsychological Status - RBANS (Randolph et al., 1998), the CDR, the MoCA and the AD8 Dementia screening (Galvin et al., 2005). Exclusion criteria common to all participants included contraindications to MRI, presence or history of neurologic disease with potential impact on cognition (e.g., Parkinson's disease), and presence or history of substance abuse.

Genetic and CSF biomarkers in PREVENT-AD subjects

In 228 PREVENT-AD subjects, DNA was isolated from 200 ul of whole blood using a QIASymphony apparatus and the DNA Blood Mini QIA Kit (Qiagen, Valencia, CA, USA). The standard QIASymphony isolation program was performed as per the manufacturer's instructions. APOE single nucleotide polymorphism (SNP) genotyping was performed using pyrosequencing (PyroMArk96) and processed with GenomeStudio (version 2010.3) using standard methods .

CSF samples were obtained by lumbar puncture in 59 subjects of the PREVENT-AD cohort. For each subject, 25 ml of CSF was centrifuged 10 minutes +/- 2000g at room temperature and aliquoted in 50 vials of 0.5 ml and frozen at -80C for further analysis. Protein levels of $A\beta_{1-42}$, total tau (tTau) and phosphorylated tau (pTau) were determined by enzyme-linked immunosorbent assay (ELISA) from Innotest technology (Fujirebio). These measurements were standardized with the European project BIOMARKAPD (Reijs et al., 2015), which intends to harmonize assays that are used to measure biological markers in neurodegenerative diseases.

MRI acquisition

The MTL-ADNI2 multisite resting-state dataset included brain imaging data acquired on 3T MRI scanners (Table 2). Vendors differed between sites (Siemens Magnetom Tim Trio in MTL sites and Phillips Achieva or Intera in ADNI2). Analyses were performed on the first usable scan, typically the baseline scan when several scans were available. Functional scan acquisition parameters varied from one site to another, notably in run duration (ranges: 5min20s-8min), number of volumes (range: 140-240 vols), voxel size (range: 3-4x3-3.6x3.3-4mm^3) and repetition time (range: 2-3s). Brain imaging data of the PREVENT-AD dataset were collected on a single 3T MRI scanner (Siemens, Magnetom Tim Trio). Two consecutive resting-state runs of 150 functional volumes were acquired, each run lasting 5min 04s. Spatial and temporal resolutions were as follows: voxel size = 4x4x4mm^3 and repetition time = 2000ms. Table 2 reports scan acquisition parameters for all data.

rs-fMRI preprocessing

Datasets were preprocessed and analyzed using the NeuroImaging Analysis Kit - NIAK - version 0.12.17 (http://niak.simexp-lab.org), under CentOS with Octave (http://gnu.octave.org) version 3.6.1 and the MINC toolkit (http://bic-mni.github.io/) version 0.3.18. Analyses were executed in parallel on the "Guillimin" supercomputer

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

(http://www.calculquebec.ca/en/resources/compute-servers/guillimin), using the pipeline system for Octave and Matlab - PSOM (Bellec et al., 2012).

Each fMRI dataset was corrected for differences in timing of slice acquisitions; a rigid-body motion was then estimated using Minctracc (Collins and Evans, 1997) for each time frame, both within and between runs, as well as between one fMRI run and the T1 scan for each subject. The T1 scan was itself non-linearly co-registered to the Montreal Neurological Institute (MNI) ICBM152 stereotaxic symmetric template (Fonov et al., 2011), using the CIVET pipeline (Ad-Dab'bagh et al., 2006). The rigid-body, fMRI-to-T1 and T1-to-stereotaxic transformations were all combined to resample the fMRI in MNI space at a 3 mm isotropic resolution. To minimize artifacts due to excessive motion, all time frames showing an average frame displacement (FD) greater than 0.5 mm were removed (Power et al., 2012). The following nuisance covariates were regressed out from the fMRI time series: slow time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (accounting for 95% variance) of the six rigid-body motion parameters and their squares (Giove et al., 2009; Lund et al., 2006). The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel. A more detailed description of the pipeline can be found on the NIAK website (http://niak.simexp-lab.org/pipe_preprocessing.html).

Individual voxel-wise connectivity maps based on large-scale network templates

For all 361 subjects included in the analyses, we computed voxel-wise connectivity maps associated with each of 7 network templates extracted from a functional brain atlas generated on 200 healthy subjects (https://doi.org/10.6084/m9.figshare.1285615.v1). The atlas included cerebellar, limbic, visual, motor, default-mode, fronto-parietal and salience networks. For each subject and each network, a network connectivity map was obtained by computing the Fisher-transformed Pearson's correlations between the average time course within the network template and the time course of every voxel in the brain grey matter. For each network, subject by voxel connectivity matrices were defined at the group level, separately for the MTL-ADNI and PREVENT-AD samples. Two general linear models were used to regress the following confounds on the group connectivity matrices: age, sex and residual (after scrubbing) FD, as well as acquisition protocols / study using dummy variables, i.e. MNI, CRIUGMa, CRIUGMb, ADNIa, ADNIb. The inclusion of constant terms in the models effectively normalized network connectivity maps to a zero grand mean across all subjects, separately for the MTL-ADNI and PREVENT-AD samples.

Network subtypes defined by a cluster analysis in MTL-ADNI2 subjects

For each of the 7 rsfMRI networks, a subject by subject similarity (Pearson's correlation) matrix summarized the between-subject correspondence of connectivity maps for all pairs of the 130 subjects in the MTL-ADNI multisite sample. A hierarchical cluster

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

analysis was performed to identify 3 clusters of subjects whose network connectivity maps were similar in terms of spatial extent and/or strength. For each cluster, we defined a subtype of functional connectivity as the average connectivity map for subjects within this cluster. In total, there were thus 21 subtypes being investigated. Subtype weights were obtained by calculating the correlation between individual connectivity maps and each of the network subtype maps. Weights thus range between -1 and 1, with 1 meaning perfect correspondence, 0 lack of correspondence and -1 perfect but inverted correspondence.

Statistical tests of association with clinical symptoms in MTL-ADNI2 subjects

We tested the association between subtypes of network connectivity and clinical symptoms in the 130 MTL-ADNI2 subjects. To this end, we employed two distinct statistical approaches: one approach treated subtypes as discrete units, where each subject belongs to one and only one cluster; a second approach used subtype weights, which are continuous measures. Despite these conceptual differences, we expected both statistical approaches to provide mostly concordant results. In the first approach, Chi2 tests were used to reveal unequal distributions of HC and ADMCI patients across the subtypes of each network. We report Cramer's V effect sizes for which values of 0.1, 0.3 and 0.5 are respectively termed small, medium and large. In our second approach, we used general linear models to test separately the associations between the weights of each network subtype and clinical symptoms (HC vs. ADMCI). Because confounds (age, sex, rFD, sites) were regressed out prior to conducting this analysis, no factors of interest were entered in the general linear model. We provide Cohen's f2 effect sizes for which values of 0.02, 0.15 and 0.35 are termed small, medium and large, respectively (Cohen, 1988). In both statistical approaches, results were deemed significant if they survived false-discovery rate (FDR) correction at $q<0.05$ across networks and subtypes.

Matching of FH subjects to PAS subtypes

We next aimed to match connectivity maps in 231 cognitively normal FH older adults with PAS subtypes identified in the MTL-ADNI2 dataset. For each network and each PREVENT-AD subject, subtype weights were obtained by correlating his/her connectivity map (averaged over 2 runs) with each of the 3 subtype maps identified in the clinical sample. Each FH subject was assigned to the subtype for which the weight was maximal. We then tested, for each network, the similarity of subject distributions across subtypes between FH subjects in the PREVENT-AD cohort vs the distribution of ADMCI patients or HC subjects in the MTL-ADNI multisite sample. Chi2 tests were used to assess significance of differences in distributions and Cramer's V values described effect sizes.

Test-retest reliability of MTL-ADNI2 subtypes in FH subjects

Intra-class correlation coefficients quantified the reproducibility of weights between the two consecutive resting-state runs of the PREVENT-AD cohort. With 7 networks and 3

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

subtypes, we thus obtained 21 ICC measures. ICC measures were interpreted as follows (Cicchetti, 1994): less than 0.40 = poor, between 0.40 and 0.59 = fair, between .60 and 0.74 = good, between 0.75 and 1 = excellent.

Statistical tests of association with AD biomarkers

We finally assessed whether the subtype weights of FH subjects would be associated with known biomarkers or risk factors of AD in PREVENT-AD. Namely, we investigated the possible association between APOE4 genotype, CSF $A\beta_{1-42}$ and Tau levels with symptom associated network subtypes. Associations were tested in the framework of general linear models and were considered significant if they survived false-discovery rate (FDR) correction at $q<0.05$ across networks and subtypes. Because confounds (age, sex, rFD) were regressed out prior to conduct this analysis, no factors of interest were entered in the general linear models. Effect sizes are reported with Cohen's f2 measures.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Supplemental Information

NA

Author contributions

Conceptualization, P.O., and P.B.; Data acquisition, M.S., C.M., A.S., A.D., J.P., P.R-N.; Analysis, P.O., A.T., M.S., C.M., C.D.; Writing - original draft, P.O., A.B., P.B.; Writing - review and editing, A.T., M.S., C.M., C.D., J.V., S.V., J.P.; Supervision, J.B., and P.B.; Funding, J.P., P.R-N., J.B., & P.B.

Acknowledgments

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 DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; 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.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing 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 coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

Data collection and sharing for the PREVENT-AD program were supported by its sponsors, McGill University, the Fonds de Research du Québec – Santé, the Douglas Hospital Research Centre and Foundation, the Government of Canada, the Canadian Foundation for Innovation, the Levesque Foundation, an unrestricted gift from Pfizer Canada Genome Quebec Innovation Center. Private sector contributions are facilitated

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

by the Development Office of the McGill University Faculty of Medicine and by the Douglas Hospital Research Centre Foundation (http://www.douglas.qc.ca/).

This work was supported by a salary award (Junior 1 scholarship) by Fonds de Recherche du Québec - Santé (FRQS) as well as funds by the Canadian Institutes of Health research (CIHR), the Alzheimer's Society of Canada and the Courtois foundation to PB. AB was supported by postdoctoral fellowships from the Canadian Alzheimer Society and CIHR, as well as by the Lemaire foundation. CD is supported by a bursary from the Lemaire foundation. We are grateful to Sylvie Belleville, Howard Chertkow, Louis Collins, Samir Das, Alain Dagher, Alan Evans, Alexandru Hanganu, Ouri Monchi, Amir Schmuel, Louise Theroux and Seqian Wang for sharing fMRI datasets and/or analytical tools.

Finally, we would like to acknowledge the participants of the PREVENT-AD cohort for dedicating their time and energy to help us collecting this data.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Tables

			MTL-ADNI2			PREVENT-AD
	MNI	**CRIUGMa**	**CRIUGMb**	**ADNI2a**	**ADNI2b**	
N controls	13	13	8	20	11	n/a
Mean age (s.d.)	67 (5.8)	71.2 (4.8)	72.6 (7.8)	75.3 (6.5)	75.9 (8.7)	n/a
Number male (%)	5 (38.5)	4 (30.8)	5 (62.5)	9 (45)	1 (9.1)	n/a
N ADMCI patients	13	13	8	20	11	n/a
N MCI patients	13	0	8	13	10	n/a
N AD dementia patients	0	13	0	7	1	n/a
Mean age (s.d.)	71.6 (8.4)	75 (7)	79.9 (6.1)	72 (7.9)	67 (5)	n/a
Number male (%)	5 (38.5)	2 (15.4)	3 (37.5)	7 (35)	7 (63.6)	n/a
N FH subjects	n/a	n/a	n/a	n/a	n/a	231
Mean age (s.d.)	n/a	n/a	n/a	n/a	n/a	64.1 (5.7)
Number male (%)	n/a	n/a	n/a	n/a	n/a	59 (25.5)
N $A\beta_{1\text{-}42}$	n/a	n/a	n/a	n/a	n/a	79
Mean $A\beta_{1\text{-}42}$ (s.d.)	n/a	n/a	n/a	n/a	n/a	1079.7 (280.9)
N ApoE4	n/a	n/a	n/a	n/a	n/a	228
N ApoE4 carriers (%)	n/a	n/a	n/a	n/a	n/a	78 (34.2)

Table 1. Demographics

Basic demographics (sample size, mean age, sex proportions) are given for the HC, ADMCI and FH groups. Levels of CSF $A\beta_{1\text{-}42}$ and proportions of ApoE4 carriers are given for FH subjects.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

			MTL-ADNI2			PREVENT-AD
	MNI	**CRIUGMa**	**CRIUGMb**	**ADNI2a**	**ADNI2b**	
Scanner manufacturer	Siemens	Siemens	Siemens	Phillips	Phillips	Siemens
Structural						
N channels	32	32	32	8	8	12
N slices	176	176	176	170	170	176
Voxel size (mm3)	1x1x1	1x1x1	1x1x1	1x1x1.2	1x1x1.2	1x1x1
Matrix size	256x256	256x256	240x256	256x256	256x256	256x256
FOV (mm2)	256	256	240/256	256	256	256?
TR (s)	2.3	2.53	2.3	6.8	6.8	2.3
TE (ms)	2.98	1.64	2.91	3.09	3.09	2.98
FA (degrees)	9	9	9	9	9	9
Functional						
N channels	32	32	32	8	8	12
N slices	38	33	33	48	48	32
Voxel size (mm3)	3.6x3.6x3.6	3x3x4	3x3x4	3.3x3.3x3.3	3.3x3.3x3.3	4x4x4
Matrix size	64x64	64x64	64x64	64x64	64x64	64x64
FOV (mm2)	230	192	192	212	212	256?
TR (s)	2	2	2	3	3	2
TE (ms)	30	30	30	30	30	30
FA (degrees)	90	90	90	80	80	90
No. volumes	160	240	240	140	140	150 (x 2)
Scan duration (min:s)	5:20	8:00	8:00	7:00	7:00	5:04 (x2)

Table 2. MRI acquisition protocols

Scan parameters are given for structural and functional data across the 5 MTL-ADNI samples as well as the PREVENT-AD dataset.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

References

Ad-Dab'bagh, Y., Lyttelton, O., Muehlboeck, J.S., Lepage, C., Einarson, D., Mok, K., Ivanov, O., Vincent, R.D., Lerch, J., Fombonne, E., Others, 2006. The CIVET image-processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research. In: Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping. Florence, Italy, p. 2266.

Badhwar, A., Tam, A., Dansereau, C., Orban, P., Hoffstaedter, F., Bellec, P., 2017. Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 8, 73–85.

Badhwar, A., Tam, A., Dansereau, C., Orban, P., Toro, R., Bellec, P., 2016. Resting-state network dysfunction in Alzheimer's disease: a systematic review and meta-analysis. Alzheimers. Dement.

Beach, T.G., Monsell, S.E., Phillips, L.E., Kukull, W., 2012. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. J. Neuropathol. Exp. Neurol. 71, 266–273.

Bellec, P., Lavoie-Courchesne, S., Dickinson, P., Lerch, J.P., Zijdenbos, A.P., Evans, A.C., 2012. The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows. Front. Neuroinform. 6, 7.

Braak, H., Braak, E., 1991. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 82, 239–259.

Brier, M.R., Thomas, J.B., Ances, B.M., 2014. Network dysfunction in Alzheimer's disease: refining the disconnection hypothesis. Brain Connect. 4, 299–311.

Buckner, R.L., Snyder, A.Z., Shannon, B.J., LaRossa, G., Sachs, R., Fotenos, A.F., Sheline, Y.I., Klunk, W.E., Mathis, C.A., Morris, J.C., Mintun, M.A., 2005. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. J. Neurosci. 25, 7709–7717.

Chételat, G., La Joie, R., Villain, N., Perrotin, A., de La Sayette, V., Eustache, F., Vandenberghe, R., 2013. Amyloid imaging in cognitively normal individuals, at-risk populations and preclinical Alzheimer's disease. Neuroimage Clin 2, 356–365.

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. 6, 284.

Cohen, J., 1988. Statistical power analysis for the behavioral sciences Lawrence Earlbaum Associates. Hillsdale, NJ 20–26.

Collins, D.L., Evans, A.C., 1997. Animal: Validation and Applications of Nonlinear Registration-Based Segmentation. Int. J. Pattern Recognit Artif Intell. 11, 1271–1294.

Dansereau, C., Benhajali, Y., Risterucci, C., Pich, E.M., Orban, P., Arnold, D., Bellec, P., 2017. Statistical power and prediction accuracy in multisite resting-state fMRI connectivity. Neuroimage 149, 220–232.

Delbeuck, X., Van der Linden, M., Collette, F., 2003. Alzheimer'disease as a disconnection syndrome? Neuropsychol. Rev. 13, 79–92.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Doan, N.T., Engvig, A., Persson, K., Alnæs, D., Kaufmann, T., Rokicki, J., Córdova-Palomera, A., Moberget, T., Brækhus, A., Barca, M.L., Engedal, K., Andreassen, O.A., Selbæk, G., Westlye, L.T., 2017. Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer's disease spectrum. Sci. Rep. 7, 45131.

Dong, A., Toledo, J.B., Honnorat, N., Doshi, J., Varol, E., Sotiras, A., Wolk, D., Trojanowski, J.Q., Davatzikos, C., Alzheimer's Disease Neuroimaging Initiative, 2017. Heterogeneity of neuroanatomical patterns in prodromal Alzheimer's disease: links to cognition, progression and biomarkers. Brain 140, 735–747.

Dubois, B., Hampel, H., Feldman, H.H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., Broich, K., Cavedo, E., Crutch, S., Dartigues, J.-F., Duyckaerts, C., Epelbaum, S., Frisoni, G.B., Gauthier, S., Genthon, R., Gouw, A.A., Habert, M.-O., Holtzman, D.M., Kivipelto, M., Lista, S., Molinuevo, J.-L., O'Bryant, S.E., Rabinovici, G.D., Rowe, C., Salloway, S., Schneider, L.S., Sperling, R., Teichmann, M., Carrillo, M.C., Cummings, J., Jack, C.R., Jr, Proceedings of the Meeting of the International Working Group (IWG) and the American Alzheimer's Association on "The Preclinical State of AD"; July 23, 2015; Washington DC, USA, 2016. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimers. Dement. 12, 292–323.

Elman, J.A., Madison, C.M., Baker, S.L., Vogel, J.W., Marks, S.M., Crowley, S., O'Neil, J.P., Jagust, W.J., 2016. Effects of Beta-Amyloid on Resting State Functional Connectivity Within and Between Networks Reflect Known Patterns of Regional Vulnerability. Cereb. Cortex 26, 695–707.

Filippini, N., MacIntosh, B.J., Hough, M.G., Goodwin, G.M., Frisoni, G.B., Smith, S.M., Matthews, P.M., Beckmann, C.F., Mackay, C.E., 2009. Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. Proceedings of the National Academy of Sciences 106, 7209–7214.

Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Brain Development Cooperative Group, 2011. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 54, 313–327.

Galvin, J.E., Roe, C.M., Powlishta, K.K., Coats, M.A., Muich, S.J., Grant, E., Miller, J.P., Storandt, M., Morris, J.C., 2005. The AD8: a brief informant interview to detect dementia. Neurology 65, 559–564.

Giove, F., Gili, T., Iacovella, V., Macaluso, E., Maraviglia, B., 2009. Images-based suppression of unwanted global signals in resting-state functional connectivity studies. Magn. Reson. Imaging 27, 1058–1064.

Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc. Natl. Acad. Sci. U. S. A. 101, 4637–4642.

Hwang, J., Kim, C.M., Jeon, S., Lee, J.M., Hong, Y.J., Roh, J.H., Lee, J.-H., Koh, J.-Y., Na, D.L., Alzheimer's Disease Neuroimaging Initiative, 2016. Prediction of Alzheimer's disease pathophysiology based on cortical thickness patterns. Alzheimers. Dement. 2, 58–67.

Hyman, B.T., Phelps, C.H., Beach, T.G., Bigio, E.H., Cairns, N.J., Carrillo, M.C., Dickson, D.W., Duyckaerts, C., Frosch, M.P., Masliah, E., Mirra, S.S., Nelson, P.T.,

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Schneider, J.A., Thal, D.R., Thies, B., Trojanowski, J.Q., Vinters, H.V., Montine, T.J., 2012. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimers. Dement. 8, 1–13.

Jiang, Y., Huang, H., Abner, E., Broster, L.S., Jicha, G.A., Schmitt, F.A., Kryscio, R., Andersen, A., Powell, D., Van Eldik, L., Gold, B.T., Nelson, P.T., Smith, C., Ding, M., 2016. Alzheimer's Biomarkers are Correlated with Brain Connectivity in Older Adults Differentially during Resting and Task States. Front. Aging Neurosci. 8, 15.

Jones, D.T., Knopman, D.S., Gunter, J.L., Graff-Radford, J., Vemuri, P., Boeve, B.F., Petersen, R.C., Weiner, M.W., Jack, C.R., Jr, Alzheimer's Disease Neuroimaging Initiative, 2016. Cascading network failure across the Alzheimer's disease spectrum. Brain 139, 547–562.

Korolev, I.O., Symonds, L.L., Bozoki, A.C., Alzheimer's Disease Neuroimaging Initiative, 2016. Predicting Progression from Mild Cognitive Impairment to Alzheimer's Dementia Using Clinical, MRI, and Plasma Biomarkers via Probabilistic Pattern Classification. PLoS One 11, e0138866.

Lam, B., Masellis, M., Freedman, M., Stuss, D.T., Black, S.E., 2013. Clinical, imaging, and pathological heterogeneity of the Alzheimer's disease syndrome. Alzheimers. Res. Ther. 5, 1.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.-L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? Neuroimage 29, 54–66.

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536.

Mufson, E.J., Malek-Ahmadi, M., Perez, S.E., Chen, K., 2016. Braak staging, plaque pathology, and APOE status in elderly persons without cognitive impairment. Neurobiol. Aging 37, 147–153.

Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H., 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. 53, 695–699.

Orban, P., Dansereau, C., Desbois, L., Mongeau-Pérusse, V., Giguère, C.-É., Nguyen, H., Mendrek, A., Stip, E., Bellec, P., 2017. Multisite generalizability of schizophrenia diagnosis classification based on functional brain connectivity. Schizophr. Res.

Orban, P., Madjar, C., Savard, M., Dansereau, C., Tam, A., Das, S., Evans, A.C., Rosa-Neto, P., Breitner, J.C.S., Bellec, P., PREVENT-AD Research Group, 2015. Test-retest resting-state fMRI in healthy elderly persons with a family history of Alzheimer's disease. Sci Data 2, 150043.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154.

Profenno, L.A., Porsteinsson, A.P., Faraone, S.V., 2010. Meta-analysis of Alzheimer's disease risk with obesity, diabetes, and related disorders. Biol. Psychiatry 67, 505–

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

512.
Randolph, C., Tierney, M.C., Mohr, E., Chase, T.N., 1998. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. J. Clin. Exp. Neuropsychol. 20, 310–319.
Reijs, B.L.R., Teunissen, C.E., Goncharenko, N., Betsou, F., Blennow, K., Baldeiras, I., Brosseron, F., Cavedo, E., Fladby, T., Froelich, L., Gabryelewicz, T., Gurvit, H., Kapaki, E., Koson, P., Kulic, L., Lehmann, S., Lewczuk, P., Lleó, A., Maetzler, W., de Mendonça, A., Miller, A.-M., Molinuevo, J.L., Mollenhauer, B., Parnetti, L., Rot, U., Schneider, A., Simonsen, A.H., Tagliavini, F., Tsolaki, M., Verbeek, M.M., Verhey, F.R.J., Zboch, M., Winblad, B., Scheltens, P., Zetterberg, H., Visser, P.J., 2015. The Central Biobank and Virtual Biobank of BIOMARKAPD: A Resource for Studies on Neurodegenerative Diseases. Front. Neurol. 6, 216.
Scheltens, N.M.E., Galindo-Garre, F., Pijnenburg, Y.A.L., van der Vlies, A.E., Smits, L.L., Koene, T., Teunissen, C.E., Barkhof, F., Wattjes, M.P., Scheltens, P., van der Flier, W.M., 2016. The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. J. Neurol. Neurosurg. Psychiatry 87, 235–243.
Seeley, W.W., Crawford, R.K., Zhou, J., Miller, B.L., Greicius, M.D., 2009. Neurodegenerative diseases target large-scale human brain networks. Neuron 62, 42–52.
Selkoe, D.J., 2002. Alzheimer's disease is a synaptic failure. Science 298, 789–791.
Sheline, Y.I., Morris, J.C., Snyder, A.Z., Price, J.L., Yan, Z., D'Angelo, G., Liu, C., Dixit, S., Benzinger, T., Fagan, A., Goate, A., Mintun, M.A., 2010. APOE4 allele disrupts resting state fMRI connectivity in the absence of amyloid plaques or decreased CSF Aβ42. J. Neurosci. 30, 17035–17040.
Sperling, R.A., Karlawish, J., Johnson, K.A., 2012. Preclinical Alzheimer disease—the challenges ahead. Nat. Rev. Neurol. 9, 54–58.
Stern, Y., 2006. Cognitive reserve and Alzheimer disease. Alzheimer Dis. Assoc. Disord. 20, S69–74.
Tampellini, D., 2015. Synaptic activity and Alzheimer's disease: a critical update. Front. Neurosci. 9, 423.
Tanzi, R.E., 2012. The genetics of Alzheimer disease. Cold Spring Harb. Perspect. Med. 2.
Vemuri, P., Jones, D.T., Jack, C.R., Jr, 2012. Resting state functional MRI in Alzheimer's Disease. Alzheimers. Res. Ther. 4, 2.
Wang, L., Brier, M.R., Snyder, A.Z., Thomas, J.B., Fagan, A.M., Xiong, C., Benzinger, T.L., Holtzman, D.M., Morris, J.C., Ances, B.M., 2013. Cerebrospinal fluid Aβ42, phosphorylated Tau181, and resting-state functional connectivity. JAMA Neurol. 70, 1242–1248.
Wang, L., Roe, C.M., Snyder, A.Z., Brier, M.R., Thomas, J.B., Xiong, C., Benzinger, T.L., Morris, J.C., Ances, B.M., 2012. Alzheimer disease family history impacts resting state functional connectivity. Ann. Neurol. 72, 571–577.
Yan, C.-G., Craddock, R.C., Zuo, X.-N., Zang, Y.-F., Milham, M.P., 2013. Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. Neuroimage 80, 246–262.

bioRxiv preprint first posted online Sep. 28, 2017; doi: http://dx.doi.org/10.1101/195164. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Zhang, X., Mormino, E.C., Sun, N., Sperling, R.A., Sabuncu, M.R., Yeo, B.T.T., Weiner, M.W., Aisen, P., Weiner, M., Petersen, R., Others, 2016. Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. Proceedings of the National Academy of Sciences 113, E6535–E6544.

Bellec, P., Benhajali, Y., Carbonell, F., Dansereau, C., Albouy, G., Pelland, M., Craddock, C., Collignon, O., Doyon, J., Stip, E., Orban, P., 2015. Impact of the resolution of brain parcels on connectome-wide association studies in fMRI. NeuroImage 123, 212–228.

Appendix C

A brain signature highly predictive of future progression to Alzheimer's dementia

Christian Dansereau, Angela Tam, AmanPreet Badhwar, Sebastian Urchs, Pierre Orban, Pedro Rosa-Neto, Pierre Bellec

arXiv:1712.08058v1 [q-bio.QM] 21 Dec 2017

A brain signature highly predictive of future progression to Alzheimer's dementia

Christian Dansereau[a,b,*], Angela Tam[c], AmanPreet Badhwar[a], Sebastian Urchs[c], Pierre Orban[a,e,f], Pedro Rosa-Neto[d], Pierre Bellec[a,b,*], for the Alzheimer's Disease Neuroimaging Initiative[**]

[a] *Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, CA*
[b] *Département d'Informatique et de recherche opérationnelle, Université de Montréal, Montréal,CA*
[c] *Integrated Program in Neuroscience, McGill University, Montréal,CA*
[d] *Douglas Mental Health institute, McGill University, Montréal,CA*
[e] *Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal, Montréal, CA*
[f] *Département de Psychiatrie, Université de Montréal, Montréal, CA*

Abstract

Early prognosis of Alzheimer's dementia is hard. Mild cognitive impairment (MCI) typically precedes Alzheimer's dementia, yet only a fraction (30%-50%) of MCI individuals will progress to dementia. Even when a prognosis of dementia is established using machine learning models and biomarkers, the fraction of MCI progressors remain limited (50%-75%). Instead of aiming at precise diagnosis in large clinical cohorts known for their heterogeneity, we propose here to identify only a subset of individuals who share a common brain signature highly predictive of oncoming dementia. This signature was discovered using a machine learning model in a reference public sample (ADNI), where the model was trained to identify patterns of brain atrophy and functional dysconnectivity commonly seen in patients suffering from dementia ($N = 24$), and not seen in cognitively normal individuals ($N = 49$). The model then recognized the same brain signature in 10 MCI individuals, out of $N = 56$, 90% of which progressed to dementia within three years. This result is a marked improvement on the state-of-the-art in prognostic precision, while the brain signature still identified 47% of all MCI progressors ($N = 19$). We thus discovered a sizable MCI subpopulation with homogeneous brain abnormalities and highly predictable clinical trajectories, which may represent an excellent recruitment target for clinical trials at the prodromal stage of Alzheimer's disease.

*Corresponding authors: christiandansereau@gmail.com, pierre.bellec@criugm.qc.ca

**Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

1. Introduction

Alzheimer's disease (AD) is the most common age-related neurodegenerative disorder. The typical progression of late-onset, sporadic AD comprises a lengthy preclinical stage, a prodromal stage of mild cognitive impairment (MCI), and a final stage of dementia. Usually, by the time patients suffer from dementia, severe and irreversible neurodegeneration has already occurred. In order to be effective, therapies should likely be initiated at earlier stages of the disease. For this reason, many works have aimed at finding biomarkers that can predict future progression to AD dementia at the prodromal or even preclinical stages (Rathore et al., 2017; Orban et al., 2017). Accurate prediction of progression from MCI to AD dementia has however proven to be challenging, likely due to the considerable heterogeneity in brain pathology underlying both of these conditions (Rathore et al., 2017). We propose here to work around the heterogeneity issue by identifying a subset of individuals with MCI who share a homogeneous brain signature highly predictive of progression to AD dementia.

A clinical diagnosis of Alzheimer's dementia is primarily established on the basis of amnestic (e.g. memory) or nonamnestic (e.g. language, visual, executive) cognitive symptoms that interfere with the patient's activities of daily living. The diagnosis also requires the absence of evidence for concomitant neurological diseases that can substantially affect cognition, such as Lewy body dementia, fronto-temporal dementia or vascular dementia (McKhann et al., 2011). MCI show a noticeable and measurable decline in cognitive abilities, including memory and thinking skills, yet this decline is not severe enough to qualify for dementia (Petersen et al., 2014). While MCI is considered an intermediate stage between the expected cognitive decline of normal aging and the more-serious decline of dementia, not all MCI patients progress to Alzheimer's dementia. Across 41 robust MCI cohort studies, an overall annual conversion rate of 6.5% to Alzheimer's dementia was reported (Mitchell 2009). A modest conversion to dementia of 30-50% even in long-term ($>$ 5 years) observational studies, highlights the heterogeneity present in the MCI population.

Imaging biomarkers and machine learning algorithms are increasingly used to complement neuropsychological testing for AD diagnosis and prognosis (Dubois et al., 2007; Rathore et al., 2017). Established imaging biomarkers of AD are Positron Emission Tomography (PET) glucose metabolism, beta-amyloid and tau deposits (Fodero-Tavoletti et al., 2011; Sperling et al., 2011), as well as non-invasive structural magnetic resonance imaging (sMRI) brain atrophy (Lerch et al., 2005). Currently the state-of-the-art performance on the most popular reference dataset, assembled by the Alzheimer's disease neuroimaging initiative (ADNI), reaches 95% accuracy to classify AD vs cognitively normal (CN) (Fan et al., 2008b; Zhu et al., 2014; Xu et al., 2015; Zu et al., 2016), and 80% accuracy to identify patients with MCI who will progress to AD dementia in the next three years (Mathotaarachchi et al., 2017; Moradi et al., 2015; Eskildsen et al., 2013; Wee et al., 2013; Gaser et al., 2013; Davatzikos et al., 2011; Koikkalainen

et al., 2011; Misra et al., 2009). Accuracy scores, however, are difficult to interpret in isolation. Korolev et al. (2016) for example, separately reported the specificity (76%, proportion of stable MCI being correctly identified), sensitivity (83%, proportion of progressor MCI being correctly identified), and precision (80%, proportion of actual progressors amongst individuals identified as such). Precision is, in other words, the rate of progression in the subpopulation of MCI patients for which the machine learning algorithm makes a prognosis of dementia. Precision is thus a key metric for enrichment in clinical trials, as it dictates how many patients will decline in the absence of treatment. For a given sensitivity and specificity, the precision depends on the baseline rate of progression in the original MCI sample. The progression rate observed in the sample used in this paper (ADNI2) is 34%, and corresponds with the range typically observed in other cohorts followed for over 3 years (Mitchell et al., 2009). Adjusted to a 34% baseline progression, the precision levels reported so far in the literature ranged from 50% to 75%, see Table 1. There is, therefore, substantial margin for improvements in terms of prognostic precision for AD dementia within 3 years, which is the focus of this work.

The precision of imaging-based diagnosis of AD in past studies is likely limited by the pathophysiological heterogeneity of clinical diagnosis. The actual cause of dementia, AD or otherwise, can currently only be confirmed by a post mortem brain examination. The hallmarks of AD are the accumulation of beta-amyloid plaques and tau protein neurofibrillary tangles in the brain, as well as marked atrophy of the medial temporal lobe. The analysis of Beach et al. (2012) revealed an important mismatch between clinical and histopathological diagnoses: sensitivity ranged from 71% to 87%, and specificity ranged from 44% to 71%, depending on the level of confidence in the clinical and pathophysiological examination. In particular, 30% of patients diagnosed with AD dementia in that study had no or minimal signs of AD pathology, while markers of AD pathology has been observed in 10% to 30% of cognitively normal (CN) individuals, as well as 40% of patients diagnosed with non-AD dementia (Beach et al., 2012). In the MCI population Petersen et al. (2014) reported prevalence of 4.8% per year. In addition to such incorrect diagnoses, co-occurrence of other age-related neurodegenerative diseases is common, including vascular brain injury, Lewy body disease, or hippocampal sclerosis Rabinovici et al. (2017); Jellinger et al. (2014). Individuals suffering from MCI in particular exhibit a wide range of brain pathologies (Stephan et al., 2012). In summary, the clinical diagnoses currently used are often incorrect (wrong underlying disease) and incomplete (missing several interacting diseases). Brain markers likely cannot be linked to clinical diagnoses with high precision in this context.

In this work, we proposed a new machine learning model that worked around the issue of heterogeneity by identifying a subgroup of patients who (1) shared homogeneous brain abnormalities; and (2) had a highly predictable clinical diagnosis or prognosis. A cluster analysis was first used on structural and functional magnetic resonance images to identify subtypes of brain atrophy and functional connectivity in a sample mixing CN individuals with patients suffering from AD dementia. Using a novel two-step procedure, a model was trained to identify

a brain signature mixing subtypes from different modalities, that was highly specific of patients with dementia. We then identified a subset of MCI patients presenting with this brain signature, and evaluated the rate of progression to dementia within 3 years in these individuals.

2. Results

Simple simulation

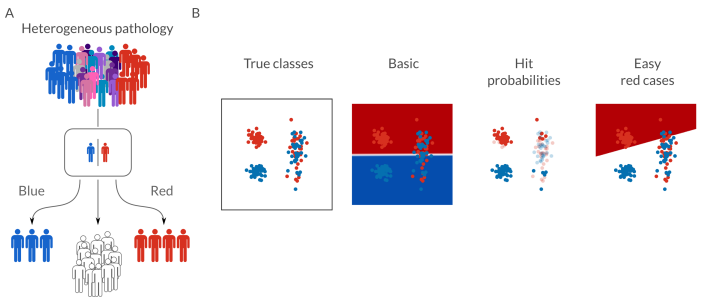

Figure 1: Panel A show the identification of easy cases for each class, Panel B prediction of clinical labels in a two-class problem, in the presence of heterogeneous labels in a subset of the data. The first column shows the initial classification problem with the distribution of the two classes. The second column shows a basic classifier decision hyperplane. The third column shows the subjects that have been flagged as high hit probability in hard color and the low hit probabilities with some transparency. The fourth column shows the final decision hyperplane of the red subjects with the HPS signature.

We first illustrated the behaviour of the proposed method with a simple simulation (Figure 1A). The task was to classify two classes using a separation line: blue dots for controls and red dots for patients. The distribution of both red and blue subjects was heterogeneous, in the sense that each distribution was a mixture of several Gaussian classes. Some of these classes were clearly separable, yet others were not, with blue and red points closely overlapping (maybe because of incorrect or incomplete diagnosis). When a standard classifier was applied on that data, it identified a separation line making a tradeoff in sensitivity and specificity across all examples (see Figure 1B, second column). By perturbing the data, it was possible to identify the "easy cases", i.e. the data point that can be reliably classified correctly: more opaque points are associated with more reliable predictions and clearly identify the two well-separated classes at the top in Figure 1B, third column. A separate model was then trained to identify the"easy cases" red points (see Figure 1B, fourth columns). The resulting prediction of red labels had limited sensitivity, as the problematic cases were not being detected at all, but it had near perfect specificity and precision.

Multimodal imaging markers

We extracted multimodal (structural and functional) measures of brain organization that could be used for automated AD diagnosis. The measures were derived from the baseline MRI scans of the ADNI2 cohort, which included anatomo-functional imaging for CN subjects (N=49) as well as patients suffering from AD dementia (N=24) (available sample size post quality-control on 10/2016). We decided to include a range of different measures previously shown to be sensitive markers of AD dementia. These included gray matter (GM) thickness (Querbes et al., 2009; Eskildsen et al., 2013), GM volume of various brain structures (Karas et al., 2004), as well as seed-based functional MRI (fMRI) connectivity maps generated for 20 intrinsic connectivity brain networks (Urchs et al., 2017).

Substantial inter-individual variations were observed in the distribution of normalized brain imaging measures. For example, some subjects showed higher- or lower-than average volumetric measures across extensive brain territories, such as the right medial occipital cortex in subject 1 (lower) and subject 73 (higher), see Figure 2A. We investigated whether such patterns could be found systematically in a subgroup of subjects. For this purpose, we quantified the similarity of GM volume maps between all pair of subjects using a Pearson correlation coefficient (Figure 2B). A cluster analysis revealed the presence of three subgroups of subjects with homogeneous GM volume maps. These subgroups were apparent as squares with high similarity values along the diagonal of the inter-subject similarity matrix, Figure 2B. These squares outline the spatial similarities of GM volume maps of subjects within a specific subgroup. By contrast, low similarity values were observed in elements outside of these squares, which corresponded to pairs of subjects falling into different subgroups. A subtype template was generated for each subgroup by averaging the maps of individuals within that subgroup, Figure 2B). In particular, subtypes 2 and 3 of GM volumetric maps reproduced the pattern observed in the occipital cortex of subjects 1 and 73, respectively. The separation between clusters was not clear-cut in matrix 2B, suggesting a continuum rather than discrete subtypes. We thus extracted a continuous measure (Pearson's correlation) of similarity, called "subtype weights", between each individual map and each subtype map, Figure 2D). The subtyping procedure outlined above was applied independently for each type of measure (volumetric, cortical thickness, rs-fMRI) and each brain network (for rs-fMRI). We confirmed by visual inspection the presence of at least three subtypes for each modality/network, which we thus selected as a common number of subtypes across all modalities/networks for subsequent analyses.

Prediction of AD

We established a baseline performance for automatic classification of CN vs AD subjects using a well established machine learning model, i.e. a linear support vector machine model (SVM) (Cortes and Vapnik, 1995). The model reached 70% precision (specificity 86%, sensitivity 67%) using tenfold cross-validation and multimodal (fMRI + sMRI) subtype weights, Figure 3. Training

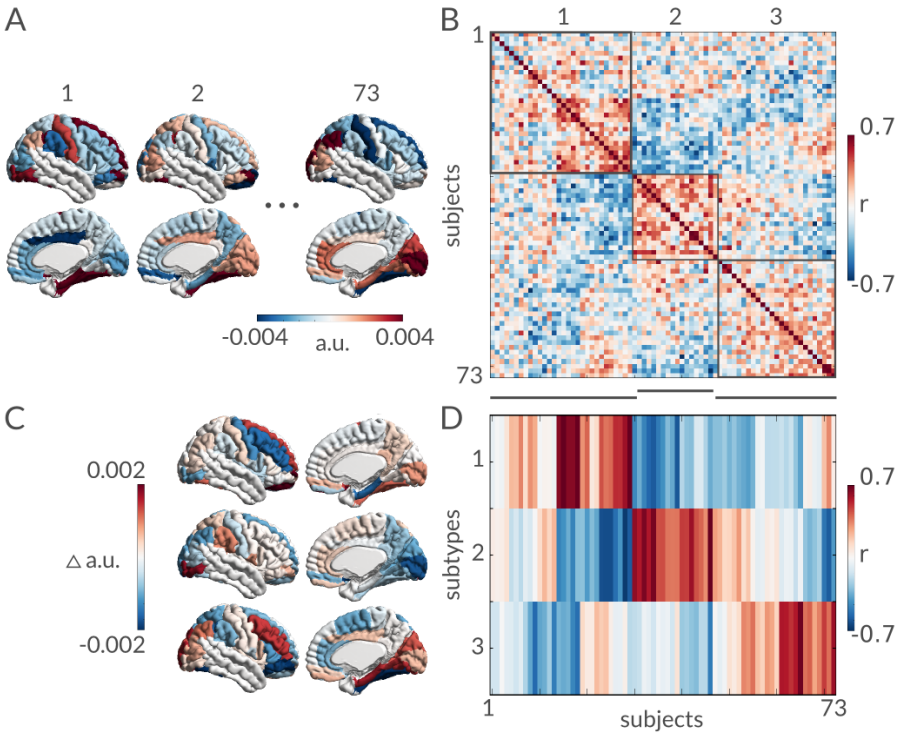

Figure 2: Demeaned gray matter volume measures of the right hemisphere. Panel A shows individual maps and the correlation of every subject with all other subjects in Panel B. Panel C shows the subtypes templates representing subgroups in the dataset. Panel D shows the association of each individual map in A with each subtype template in C.

only on fMRI subtypes or only on sMRI subtypes yielded lower performances: 38% precision (specificity 47% and sensitivity 67%) for fMRI alone and 67% precision (specificity 84%, sensitivity 67%) for sMRI alone. Note that, during cross-validation, the training of the model included both the generation of subtypes and the optimization of the SVM parameters.

Identifying easy cases

As we outlined in the introduction, the core idea of our approach was to identify a subset of subjects for which clinical labels are easy to predict, such as the points on the left in Figure 1A. To identify these "easy cases", we randomly perturbed the input data of the SVM model many times through subsampling, and assessed the hit probability for any given subject to be properly classified. We found that 68% of individuals had a perfect (100%) hit probability, with a small subset of subjects (18%) exhibiting less reliable predictions (hit-probability $< 90\%$), Supplementary material S1). We defined the "easy cases" as the subgroup of individuals reaching perfect hit probability.

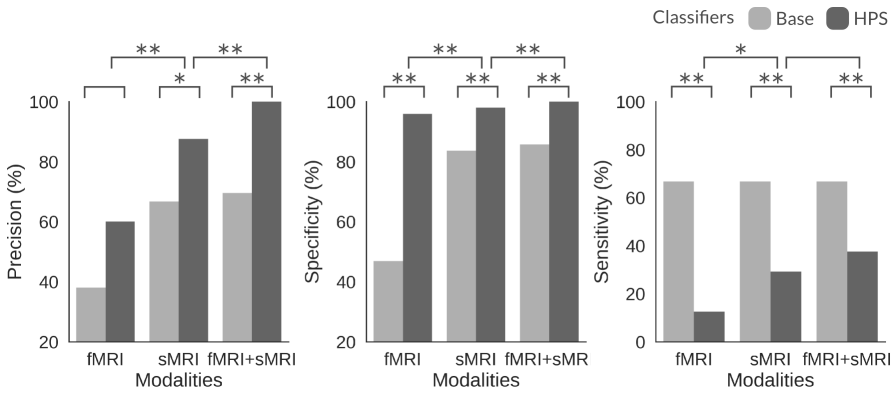

Figure 3: Figure shows the precision, specificity and sensitivity of the three modalities (fMRI, sMRI and fMRI+sMRI) at each stage (Base: basic classifier and HPS: highly predictive signature). Significant differences are shown with $*$ for $p < 0.05$ and $**$ for $p < 0.001$).

Predicting easy cases

The next step of the method was to train a logistic regression (Fan et al., 2008a) to predict the AD "easy cases" Figure 6B, analogous to the rightmost column of Figure 1B. The full multi-stage process of subtype extraction, hit probability estimation, and logistic regression was cross-validated using a ten-fold scheme in order to generate the performance of the prediction of AD "easy cases". A perfect 100% precision (specificity 100%, sensitivity 36%) was reached for AD "easy cases", using multimodal structural and functional features. The multimodal HPS performance was a significant improvement (in precision and specificity, $p < 0.001$) compared to the model trained on fMRI only, precision of 60% (specificity 96%, sensitivity 13%), and sMRI only, precision of 88% (specificity 98%, sensitivity 29%), see Figure 3. Compared to the reference SVM model, with multimodal features, the precision of our proposed HPS model was improved by a wide margin (30%, $p < 0.001$), as well as the specificity (15%, $p < 0.001$), at the cost of a marked loss in sensitivity (30%, $p < 0.001$). See Supplementary material Table S2 for a list of the performance of each model.

Highly predictive brain signature

The logistic regression model used to predict AD "easy cases" is based on a set of coefficients, which give more or less weight to a particular subtype and modality. As such, the individuals flagged as AD "easy cases" can be seen as sharing a brain HPS, composed of a combination of subtype maps. The logistic model may in theory ignore a subtype or a modality entirely, by setting the corresponding weights to zero. In practice, we found that the HPS relied on all three types of measures (functional connectivity, cortical thickness, and gray matter volume), Figure 4A. To rank the contribution of each modality in the decision process, we computed the absolute sum of the coefficients for

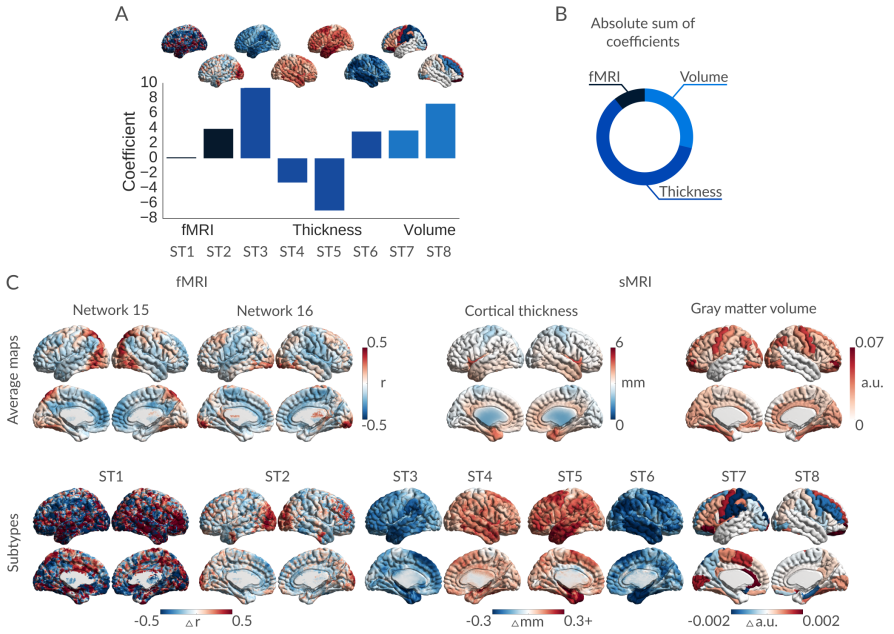

Figure 4: Panel A shows the contribution of each modality to the decision, the ratios are computed by the sum of the absolute coefficient for each modality. Panel B shows the coefficients of the high-confidence prediction model for each subtype map. Panel C shows, on top, the average maps for each modality and on the bottom the subtype maps used for the high-confidence prediction.

each measure, relative to the sum of all absolute coefficients (Figure 4B). The thickness was the most important measure (60%), followed by the volumetric measures (29%), and finally functional connectivity (11%). The highest contributions came from four subtypes of thickness: bilateral patterns of cortical atrophy in temporal, sagittal and frontal areas (one subtype per hemisphere), and bilateral, opposite patterns of increased thickness (one subtype per hemisphere), Figure 4C. Two lateralized volumetric subtypes showed gray matter volume loss in the left motor, and right frontal areas as well as a gray volume increase in the left frontal and limbic regions. Finally, one functional subtype was very noisy and barely contributed to the model, while the other highlighted a connectivity subtype connecting the visual network with frontal areas.

Prediction of progression to dementia

We applied the HPS model to patients with MCI from the ADNI2 cohort, with the hypothesis that those with the signature would likely progress to AD dementia. The imaging sample for this experiment included the baseline structural and functional scans of all MCI patients in the ADNI2 cohort ($N = 79$). We further stratified the patients with MCI into stable MCI (sMCI, $N = 37$), i.e. most recent clinical status remains MCI with at least 36 months follow up, and

Table 1: Supervised classification of MCI progression to AD dementia using the ADNI database. Progression time was establish if the the subject progresses to AD status in the next 36 months. Significant improvement of our method compared to each paper for the adjusted accuracy and precision (adjusted for a pMCI ratio of 34% comparable to our sample) and specificity are shown with $*$ for $p < 0.05$ and $**$ for $p < 0.001$) and conversely significant decrease in sensitivity of our method compared to each paper. Adjusted accuracy (Acc), adjusted precision (Prec), specificity (Spec), sensitivity (Sens)

	N	Acc	Prec	Spec	Sens
Article	sMCI/pMCI		adjusted		
Dansereau et al. (This paper)	37/19	80%	**90%**	**97%**	47%
Mathotaarachchi et al. (2017)	230/43	82%	74%	87%*	71%*
Korolev et al. (2016)	120/139	79%	65%*	76%**	83%*
Moradi et al. (2015)	100/164	78%	63%*	74%**	87%**
Eskildsen et al. (2013)	134/149	67%**	52%**	68%**	66%
Wee et al. (2013)	111/89	77%	68%*	84%**	64%
Gaser et al. (2013)	62/133	80%	70%*	84%**	71%*
Davatzikos et al. (2011)	170/69	79%	63%*	71%**	**95%****
Koikkalainen et al. (2011)	215/154	73%*	58%*	71%**	77%*
Misra et al. (2009)	76/27	67%**	51%**	60%**	80%*

progressors (pMCI, $N = 19$), i.e. individuals whose most recent known clinical status is AD dementia, with progression from MCI to AD dementia occurring within 37 months. The HPS model selected a subset of 10 MCI subjects. Using the longitudinal follow-up clinical data provided by ADNI2, we found that 9 out of 10 of these subjects were pMCI (precision of 90%, specificity of 97%, sensitivity of 47%), compared to 34% pMCI in the whole MCI sample ($p < 0.001$), Figure 5A. Within the HPS subgroup, the time to progression from baseline to the first evaluation of AD dementia appeared uniformly distributed from 5 to 37 months, with 50% subjects progressing after 24 months (Figure 5C). In addition, 100% of the MCI participants flagged as HPS were tested positive for beta amyloid deposition with AV45 testing, compared to a 69% rate in the whole MCI sample ($p < 0.05$), Figure 5A. The rate of ApoE4 carriers in the HPS subsample was 78%, compared to 55% in the whole MCI group ($p > 0.05$), Figure 5A. A similar observation could be made regarding the rate of male of 70% in the HPS subsample and 52% in the whole MCI group ($p > 0.05$). Finally the average age in the HPS group was of 74 years ± 7 and 71 years ± 7 for the whole MCI group ($p > 0.05$).

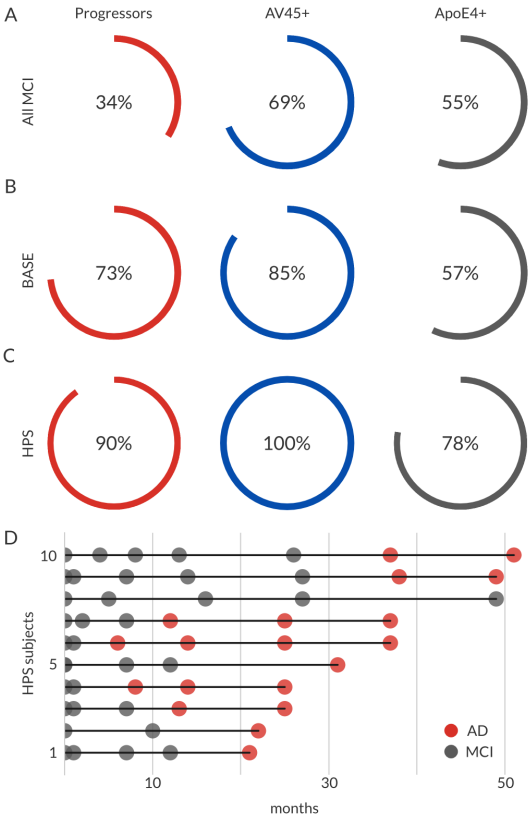

Figure 5: Statistics on the MCI showing the signature. Panel A shows the percentage of MCI who progress to AD, the percentage of subjects positive for beta amyloid deposits using the AV45 marker and the percentage of carriers of one or two copies of the ApoE4 allele for the entire MCI cohort. Panel B shows the same statistics for the selection of the base classifier while Panel C displays statistics for subjects flagged as HPS. Panel D shows the clinical status of each HPS subject over time from the baseline scan.

3. Discussion

The main goal of this work was to develop an imaging-based AD diagnosis model with high precision and specificity. The proposed HPS approach reached excellent performance in these respects: 100% precision and specificity when distinguishing patients with AD dementia from CN participants and 90% precision, 97% specificity when predicting which MCI patients will progress to dementia, up to three years prior to onset (see Table S2). These results represent a substantial improvement in precision over the state-of-the-art on this task, see Table 1. No data from MCI patients were used to train the model, which removes the possibility of a bias due to improper cross validation. The only HPS subject with MCI improperly classified as a progressor had a series of four notes attached to his visits in the ADNI database, reporting declining cognitive performance at each visit, with a marked decline at the last visit. This decline was not severe enough for a diagnosis of AD dementia. The subject had no follow up available after 36 months, for unknown reasons.

The high specificity of the HPS model came at the cost of a limited sensitivity: 38% when distinguishing patients with AD dementia from CN participants, and 47% when predicting which MCI patients would progress to dementia, which is significantly less than most recent published models, see Table 1. The HPS model is not designed to be sensitive, as it is trained to recognize a particular, homogeneous brain signature present in only a fraction of the participants. The results of (Beach et al., 2012) suggest that only about half of patients diagnosed with AD dementia have clear AD brain markers post-mortem. The observed sensitivity of 38%- 47% is thus consistent with the idea that the HPS model is picking on a typical brain presentation of AD that is already present at the prodromal stage of the disease. Note that there was no need for patients with MCI to have the same degree of atrophy as patients with AD dementia to be recognized as HPS, as long as these patients presented with a similar spatial distribution of the atrophy, relative to other brain regions.

The anatomical features selected by the method were in line with recent subtyping works, e.g. (Hwang et al., 2015), showing predominant atrophy in the temporal lobe, as well as the temporo-parietal juncture, in particular. The functional maps were more difficult to interpret, and seemed to capture some noise pattern. They still made a significant improvement in the performance of the HPS model. Because of the regularization in the logistic regression used to build the HPS model, features coming from different modalities did compete to be selected in the model. If redundant features existed, the ones with largest predictive power were selected by the classifier. This may explain why the selected functional subtypes did not involve the regions showing atrophy in the structural subtypes. We hypothesized that the HPS inferred from the AD vs CN prediction would also be useful to predict if a subject at the prodromal stage (MCI) would progress to dementia. Our results did validate this logic, but alternative strategies may be investigated in the future, e.g. training a model directly on the progressor vs stable MCI.

A limitation of the present study was a moderate sample size, with $N = 56$

patients suffering from MCI. While the ADNI database is large, resting-state fMRI has only been added to the protocol in the later stages of the study, ADNI GO and ADNI2. In addition, fMRI was only acquired on a third of the participants, even after it was added to the protocol. Because of the early role of synaptic dysfunction in AD, and the potential ability of fMRI to capture such dysfunction, we wanted to build an anatomo-functional diagnostic tool. But this choice did limit the sample size of our study since the selected subjects needed to have imaging data of the two modalities and pass their respective quality control assessment. Even with a larger sample size, another limitation of the ADNI dataset is that it does not reflect the diversity of cases observed in real-life clinical practice. Participants were in particular screened to exclude vascular dysfunction, which is a common comorbidity in AD (Gorelick et al., 2011). Resources with more inclusive enrollment criteria will help to better assess the generalizability of a biomarker-based AD diagnosis.

The most direct application of the HPS model is its use for population enrichment in pharmaceutical clinical trials on AD (Woo et al., 2017; Mathotaarachchi et al., 2017). By recruiting patients who would normally progress to AD dementia, such an enrichment would increase the effect size of the drug while reducing the sample size needed to demonstrate efficacy and therefore would also reduce the cost of the trial. The HPS brain signature is not shared among all the AD dementia population (making it a subtype), but is common enough to represent a substantial portion of participants of interest (about a third of AD dementia subjects and half of MCI progressors). An alternative enrichment strategy, more geared towards generalizability, would be to only exclude subjects that will very likely not progress to AD dementia. The HPS method thus brings us closer to precision medicine by proposing a middle ground between traditional clinical cohorts and an entirely individual medicine.

In this manuscript, we focused exclusively on two MRI modalities. Our rationale was that MRI is non-invasive and already widely used in patient care of elderly populations. Beta amyloid and tau PET imaging, by contrast, are more expensive and less available, while lumbar punctures are invasive. Nevertheless, as shown in our results the combination of multimodal factors may help to improve precision, specificity and sensitivity. Since the sensitivity of each modality to abnormality may vary across the disease stages it may be beneficial to combine them to obtain complementary information. It will therefore be important in the future to see if a combination of PET imaging, blood tests targeting specific markers, cognitive scores, genetic factors, lifestyle factors, or others can help create stronger or multiple HPS that would in effect increase the sensitivity of the model at earlier stages of the of Alzheimer's disease.

4. Acknowledgments

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 DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National

Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; 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.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health [1]. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The computational resources used to perform the data analysis were provided by Compute Canada[2]. This project was funded by NSERC grant number RN000028 and the Canadian Consortium on Neurodegeneration in Aging (CCNA), through a grant from the Canadian Institute of Health Research and funding from several partners including SANOFI-ADVENTIS R&D. CD is supported by a salary award from the Lemaire foundation and Courtois foundation. PB is supported by a salary award from "Fonds de recherche du Québec – Santé" and the Courtois foundation.

5. Materials and methods

Dataset

All functional and structural data were obtained from the Alzheimer's Disease Neuroimaging Initiative 2 (ADNI2) sample, a longitudinal standardized acquisition including three populations: cognitively normal subjects ($N = 49$, 46% male, 74±7 years of age), patients with mild cognitive impairment ($N = 56$, 51% male, 72±7.5 years of age) and patients with dementia due to AD ($N = 24$, 46% male, 72 ± 7 years of age). All participants gave their written informed consent to participate in the ADNI2 study, which was approved by the local ethics committee of participating institutions across North America. The consent form included data sharing with collaborators as well as secondary analysis. The present secondary analysis of the ADNI2 sample was approved by the local ethics committee at CRIUGM, University of Montreal, QC, Canada. All resting-state fMRI and structural scans were acquired on 3T Philips scanners with 8

[1] `www.fnih.org`
[2] `https://computecanada.org/`

channel head coils. We performed analyses on the first usable scan (typically the baseline scan) from ADNI2.

The acquisition parameters were as follows: structural scan 170 slices, voxel size 1x1x1.2 mm3, matrix size 256x256, FOV 256 mm2, TR 6.8 s, TE 3.09 ms, FA 9 degrees. A functional scan of 7 min, 48 slices, voxel size 3.3x3.3x3.3 mm3, matrix size 64x64, FOV 212 mm2, TR 3 s, TE 30 ms, FA 80 degrees, No. volumes 140. For detailed information on the acquisition, see www.adni-info.org.

Extraction of functional features

Each fMRI dataset was corrected for slice timing; a rigid-body motion was then estimated for each time frame, both within and between runs, as well as between one fMRI run and the T1 scan for each subject (Collins et al., 1994). The T1 scan was itself non-linearly co-registered to the Montreal Neurological Institute (MNI) ICBM152 stereotaxic symmetric template (Fonov et al., 2011), using the CIVET pipeline (Ad-Dab'bagh et al., 2006a). The rigid-body, fMRI-to-T1 and T1-to-stereotaxic transformations were all combined to resample the fMRI in MNI space at a 3 mm isotropic resolution. To minimize artifacts due to excessive motion, all time frames showing a frame displacement, as defined in Power et al. (2012), greater than 0.5 mm were removed. An average residual frame displacement was also estimated after scrubbing for further group analyses. A minimum of 50 unscrubbed volumes per run was required for further analysis (13 subjects were rejected). The following nuisance covariates were regressed out from fMRI time series: slow time drifts (basis of discrete cosines with a 0.01 Hz highpass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (accounting for 95% variance) of the six rigid-body motion parameters and their squares (Giove et al., 2009; Lund et al., 2006). The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel. Datasets were preprocessed and analyzed using the NeuroImaging Analysis Kit - NIAK - version 0.12.17 (http://niak.simexp-lab.org), under CentOS with Octave (http://gnu.octave.org) version 3.6.1 and the MINC toolkit (http://bic-mni.github.io/) version 0.3.18. Preprocessing of MRI data was executed in parallel on the Guillimin supercomputer (http://www.calculquebec.ca/en/resources/compute-servers/guillimin), using the pipeline system for Octave and Matlab - PSOM (Bellec et al., 2012). Seed-based fMRI connectivity maps were obtained using a functional brain template of 20 networks covering the entire brain (Urchs et al., 2017). The Pearson's correlation between the average time series of each network and every voxel of the brain was computed to derive one functional connectivity map per network.

Extraction of structural features

Native individual T1-weighted MRI scans were corrected for non-uniformity artifacts with the N3 algorithm (Sled et al., 1998). The corrected volumes were then masked for brain tissues (Smith, 2002) and registered into stereotaxic space

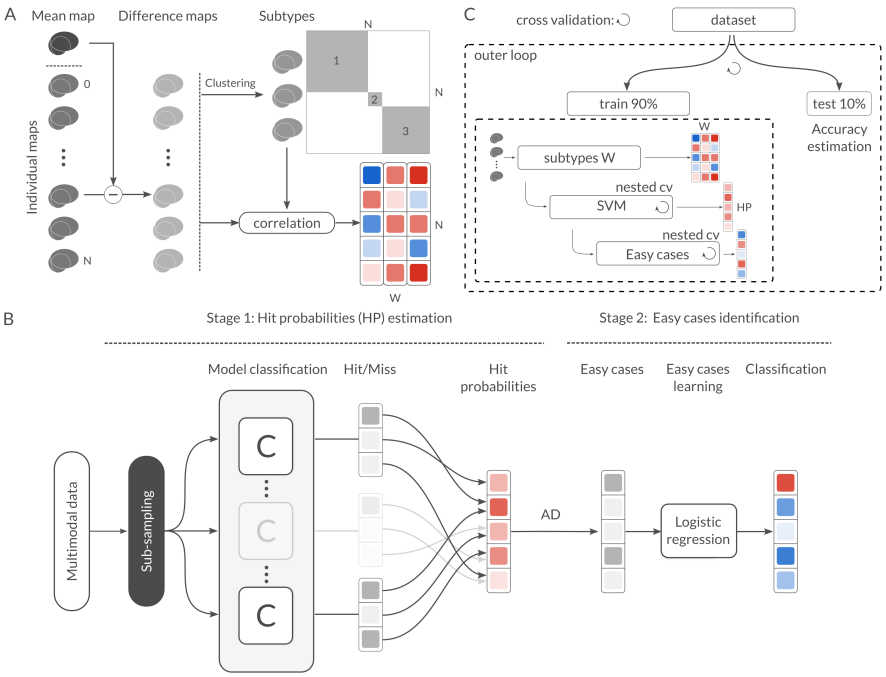

Figure 6: Panel A shows the feature extraction method called subtypes weights, Panel B framework workflow: stage 1 shows the hit probability computation based on random sub-sampling and stage 2 shows the training of dedicated classifier for each high-confidence signature. Panel C shows the nested cross-validation scheme used in this method.

(Collins et al., 1994). The registered images were segmented into gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and background using a neural net classifier (Tohka et al., 2004). The WM and GM surfaces were extracted using the Constrained Laplacian-based Automated Segmentation with Proximities algorithm (Kim et al., 2005; MacDonald et al., 2000) and were resampled to a stereotaxic surface template to provide vertex based measures and lobar segmentation (Lyttelton et al., 2007). Cortical thickness was measured in native space using the linked distance between the two surfaces across 81,924 vertices (Im et al., 2008). Surface-based cortical thickness, as well as regional volume measures, were obtained using the structural MRI images processed using the CIVET 1.1.12 pipeline for each hemisphere as described in Ad-Dab'bagh et al. (2006b). The AAL template was applied on each hemisphere (40 regions per hemisphere) to extract the regional volumetric measures. The processing pipeline was executed on the Canadian Brain Imaging Network (CBRAIN) platform, a network of five imaging centers and eight High-Performance Computers for collaborative sharing and distributed processing of large MRI databases (Frisoni et al., 2011).

Multimodal imaging subtypes

We extracted subtypes that characterize the interindividual variability within the sample comprising CN and AD participants (at the time of scanning), independently for each type of measure (functional maps, cortical thickness maps, and volumetric maps). In order to reduce the impact of some factors of no interest that may influence the clustering procedure, we regressed out the age, sex, and average post-scrubbing frame displacement from individual map, using a mass univariate linear regression model at each voxel. For each type of brain measure, we derived a spatial Pearson's correlation coefficient between all pairs of individual maps. This defined a subject x subject similarity matrix (of size 73 x 73), which was entered into a Ward hierarchical clustering procedure, as implemented in SciPy version 0.18.1 (Jones et al., 2001–; Walt et al., 2011). We selected three subgroups for each type of measure, based on a visual examination of the similarity matrix. For each type of measure, the average map of each subgroup defined a subtype. For each individual, we computed the spatial correlation of their map with each subtype. The resulting weight measures formed a matrix of size (number of subjects) x (number of subtypes), which was used as the feature space for all predictive models throughout the rest of this work.

Prediction of AD

The baseline prediction accuracy was obtained by training a SVM model with a linear kernel, as implemented in Scikit-learn Pedregosa et al. (2011) version 0.18. A tenfold cross-validation loop was used to estimate the performance of the trained model, including the entire subtyping procedure and regression of confounds. Classes were balanced inversely proportional to class frequencies in the input data for the training. A nested cross-validation loop was used (stratified shuffle split (50 splits, 20% test size)) for the grid search of the hyperparameter C (grid was 10^{-2} to 10^{1} with 15 steps). Note that the C parameter

controlled how many misclassified examples the model will tolerate by adjusting the margin size. The model was evaluated using fMRI features only, sMRI features only, and the combination of fMRI and sMRI features.

Identifying easy cases

We randomly selected subsamples of the dataset (retaining 80% of participants in each subsample) to replicate the SVM training 100 times. For each 80% subsample, a separate SVM model was trained to predict the clinical labels (CN or AD), see Figure 6B. Note that the optimal C parameter was estimated once using the whole available sample, as described above, and used across all subsamples. This was done to avoid creating major uncontrolled algorithmic variations. The linear discriminating weights of the SVM were still optimized independently for each subsample. Predictions of clinical labels were then made on the remaining 20% of subjects, that were not used for training. For a given individual, the hit probability was calculated as the frequency of correct clinical classification across all available SVM replications where the test set included that individual. Easy cases were defined as individuals with 100% hit probability.

Predicting easy cases

We trained a logistic regression classifier Fan et al. (2008a) to predict the AD easy cases. The logistic regression was trained using a L1 regularization on the coefficients, see Figure 6B. Class weight was balanced inversely proportional to class frequencies in the input data. A stratified shuffle split (100 splits, 20% test size) was used to estimate the performance of the model for the grid search of the hyper-parameter C (grid was $10^{-0.2}$ to 10^{1} with 15 equal steps). In this case, the C parameter controlled the sparseness of the weights.

Cross-validation

A nested cross-validation was performed for accuracy estimation and parameters optimization. The outer loop used to estimate the generalizability of the framework was a ten-fold cross-validation scheme. Each training fold included the full multi-stage process of subtype extraction, SVM prediction of clinical labels, identification of HPS and prediction of HPS with logistic regression. Sensitivity (true positive rate, TP), specificity (true negative rate, TN) and precision ($TP/(TP + (1 - TN))$) of the diagnosis were estimated across all test folds, in the AD vs CN prediction. Cross-validation nested inside the outer loop was used to search for the optimal hyper-parameters, Figure 6C.

Prediction of progression to dementia

The HPS was obtained by applying the subtyping and easy cases recognition to the whole CN and AD sample, and considering all subtypes associated with non-zero weights by the sparse logistic regression model in Figure 6C stage 2. The logistic regression trained on AD vs CN was used to identify MCI patients who have a HPS of AD dementia. The imaging sample for this experiment

included the baseline structural and functional scans of all patients with MCI in the ADNI2 cohort, with at least 36 months of follow-up ($N = 56$). We further stratified the patients with MCI into stable MCI (sMCI, $N = 37$)), i.e. latest clinical status is MCI, and progressors (pMCI, $N = 19$), i.e. individuals whose most recent known clinical status is AD dementia, with progression from MCI to AD dementia occurring within 36 months. Note that no AV45 imaging data or genetic data, nor any data from the MCI cohort, were used to build the HPS model.

Statistical test of differences in model performance

We generated a confidence interval on the performance (i.e. precision, specificity and sensitivity) of a given model using a Monte-Carlo simulation. Taking the observed sensitivity and specificity, and using similar sample size to our experiment, we replicated the number of true and false positive detection 100000 times using independent Bernoulli variables, and derived replications of precision, specificity and sensitivity. By comparing these replications to the sensitivity, specificity and precision observed in other models, we estimated a p-value for differences in model performance (Phipson and Smyth, 2010). A p-value smaller than 0.05 was interpreted as evidence of a significant difference in performance, and 0.001 as a strong evidence. This approach was first used in Figure 3 to contrast the performance of the HPS model to the baseline (SVM) model, both for AD vs CN and MCI progressor vs stable, as well as contrasting the performance of multimodal (fMRI+sMRI) model vs models using only fMRI or sMRI features. The same approach was used to contrast our proposed model for MCI progressor vs stable with results from the literature, in Table 1. Note that, based on our hypotheses regarding the behaviour of the HPS model, the tests were one-sided for increase in specificity and precision, and one-sided for decrease in sensitivity.

Statistical test of enrichment

The HPS model was used to select a subset of the MCI population. We tested statistically if this subgroup was enriched for (1) progression to dementia; (2) AV45+, and; (3) ApoE4+. We implemented for this purpose a Monte-Carlo simulation, where we selected 100000 random subgroups out of the original MCI sample. By comparing the proportion of progressors (respectively AV45+ and ApoE4+) in these null replications to the actual observed values in the HPS subgroup, we estimated a p-value (Phipson and Smyth, 2010) (one sided for increase). A p-value smaller than 0.05 was interpreted as evidence of a significant enrichment, and 0.001 as a strong evidence.

Public code and data

The code used in this experiment is available on a GitHub repository at the following URL[3]. An IPython Notebook is also provided with all of the figure

[3] `https://github.com/simexp/hpc`

generation scripts. Scikit-learn Pedregosa et al. (2011) version 0.18 was used for most of the machine learning algorithms and Nilearn Abraham et al. (2014) version 0.2.6 for visualization purposes. S

ADNI dataset

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics 8, 14.
URL `http://journal.frontiersin.org/article/10.3389/fninf.2014.00014`

Ad-Dab'bagh, Y., Einarson, D., Lyttelton, O., Muehlboeck, J. S., Mok, K., Ivanov, O., Vincent, R. D., Lepage, C., Lerch, J., Fombonne, E., Evans, A. C., 2006a. The CIVET Image-Processing environment: A fully automated comprehensive pipeline for anatomical neuroimaging research. In: Corbetta, M. (Ed.), Proceedings of the 12th Annual Meeting of the Human Brain Mapping Organization. Neuroimage, Florence, Italy.

Ad-Dab'bagh, Y., Einarson, D., Lyttelton, O., Muehlboeck, J. S., Mok, K., Ivanov, O., Vincent, R. D., Lepage, C., Lerch, J., Fombonne, E., Evans, A. C., 2006b. The CIVET Image-Processing Environment: A Fully Automated Comprehensive Pipeline for Anatomical Neuroimaging Research. In: Corbetta, M. (Ed.), Proceedings of the 12th Annual Meeting of the Human Brain Mapping Organization. Neuroimage, Florence, Italy.

Beach, T. G., Monsell, S. E., Phillips, L. E., Kukull, W., 2012. Accuracy of the clinical diagnosis of alzheimer disease at national institute on aging alzheimer disease centers, 2005–2010. Journal of neuropathology & experimental neurology 71 (4), 266–273.

Bellec, P., Lavoie-Courchesne, S., Dickinson, P., Lerch, J. P., Zijdenbos, A. P., Evans, A. C., 2012. The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows. Frontiers in neuroinformatics 6.
URL `http://dx.doi.org/10.3389/fninf.2012.00007`

Collins, D. L., Neelin, P., Peters, T. M., Evans, A. C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. Journal of computer assisted tomography 18 (2), 192–205.
URL http://view.ncbi.nlm.nih.gov/pubmed/8126267

Cortes, C., Vapnik, V., Sep. 1995. Support-vector networks. Machine Learning 20 (3), 273–297.
URL http://dx.doi.org/10.1007/BF00994018

Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., Trojanowski, J. Q., 2011. Prediction of mci to ad conversion, via mri, csf biomarkers, and pattern classification. Neurobiology of aging 32 (12), 2322–e19.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., Scheltens, P., Aug. 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet neurology 6 (8), 734–746.
URL http://dx.doi.org/10.1016/S1474-4422(07)70178-3

Eskildsen, S. F., Coupé, P., Garc'ıa-Lorenzo, D., Fonov, V., Pruessner, J. C., Collins, D. L., Jan. 2013. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. NeuroImage 65, 511–521.
URL http://dx.doi.org/10.1016/j.neuroimage.2012.09.058

Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., Lin, C.-J., 2008a. Liblinear: A library for large linear classification. The Journal of Machine Learning Research 9, 1871–1874.

Fan, Y., Batmanghelich, N., Clark, C. M., Davatzikos, C., Initiative, A. D. N., et al., 2008b. Spatial patterns of brain atrophy in mci patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. Neuroimage 39 (4), 1731–1743.

Fodero-Tavoletti, M. T., Okamura, N., Furumoto, S., Mulligan, R. S., Connor, A. R., McLean, C. A., Cao, D., Rigopoulos, A., Cartwright, G. A., OKeefe, G., et al., 2011. 18f-thk523: a novel in vivo tau imaging ligand for alzheimers disease. Brain 134 (4), 1089–1100.

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., Brain Development Cooperative Group, Jan. 2011. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 54 (1), 313–327.
URL http://dx.doi.org/10.1016/j.neuroimage.2010.07.033

Frisoni, G. B., Redolfi, A., Manset, D., Rousseau, M.-E., Toga, A., Evans, A. C., Jul. 2011. Virtual imaging laboratories for marker discovery in neurodegenerative diseases. Nature Reviews Neurology 7 (8), 429–438.
URL http://dx.doi.org/10.1038/nrneurol.2011.99

Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H., Initiative, A. D. N., et al., 2013. Brainage in mild cognitive impaired patients: predicting the conversion to alzheimers disease. PloS one 8 (6), e67346.

Giove, F., Gili, T., Iacovella, V., Macaluso, E., Maraviglia, B., Oct. 2009. Images-based suppression of unwanted global signals in resting-state functional connectivity studies. Magnetic resonance imaging 27 (8), 1058–1064.
URL http://dx.doi.org/10.1016/j.mri.2009.06.004

Gorelick, P. B., Scuteri, A., Black, S. E., DeCarli, C., Greenberg, S. M., Iadecola, C., Launer, L. J., Laurent, S., Lopez, O. L., Nyenhuis, D., et al., 2011. Vascular contributions to cognitive impairment and dementia. Stroke 42 (9), 2672–2713.

Hwang, J., Kim, C. M., Jeon, S., Lee, J. M., Hong, Y. J., Roh, J. H., Lee, J.-H., 2015. Prediction of alzheimers disease pathophysiology based on cortical thickness patterns. Alzheimer's & Dementia: The Journal of the Alzheimer's Association 11 (7), P541.

Im, K., Lee, J.-M., Lyttelton, O., Kim, S. H., Evans, A. C., Kim, S. I., 2008. Brain size and cortical structure in the adult human brain. Cerebral Cortex 18 (9), 2181–2191.

Jellinger, K. A., et al., 2014. Neuropathology of dementia disorders. J. Alz. Dis. Parkinsonism 4, 135.

Jones, E., Oliphant, T., Peterson, P., et al., 2001–. SciPy: Open source scientific tools for Python. [Online; accessed ¡today¿].
URL http://www.scipy.org/

Karas, G., Scheltens, P., Rombouts, S., Visser, P., Van Schijndel, R., Fox, N., Barkhof, F., 2004. Global and local gray matter loss in mild cognitive impairment and alzheimer's disease. Neuroimage 23 (2), 708–716.

Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J. M., Kim, S. I., Evans, A. C., 2005. Automated 3-d extraction and evaluation of the inner and outer cortical surfaces using a laplacian map and partial volume effect classification. Neuroimage 27 (1), 210–221.

Koikkalainen, J., Lötjönen, J., Thurfjell, L., Rueckert, D., Waldemar, G., Soininen, H., Initiative, A. D. N., et al., 2011. Multi-template tensor-based morphometry: application to analysis of alzheimer's disease. NeuroImage 56 (3), 1134–1144.

Korolev, I. O., Symonds, L. L., Bozoki, A. C., Initiative, A. D. N., et al., 2016. Predicting progression from mild cognitive impairment to alzheimer's dementia using clinical, mri, and plasma biomarkers via probabilistic pattern classification. PloS one 11 (2), e0138866.

Lerch, J. P., Pruessner, J. C., Zijdenbos, A., Hampel, H., Teipel, S. J., Evans, A. C., Jul. 2005. Focal Decline of Cortical Thickness in Alzheimer's Disease Identified by Computational Neuroanatomy. Cerebral Cortex 15 (7), 995–1001.
URL http://dx.doi.org/10.1093/cercor/bhh200

Lund, T. E., Madsen, K. H., Sidaros, K., Luo, W.-L., Nichols, T. E., Jan. 2006. Non-white noise in fMRI: does modelling have an impact? NeuroImage 29 (1), 54–66.
URL http://dx.doi.org/10.1016/j.neuroimage.2005.07.005

Lyttelton, O., Boucher, M., Robbins, S., Evans, A., 2007. An unbiased iterative group registration template for cortical surface analysis. Neuroimage 34 (4), 1535–1544.

MacDonald, D., Kabani, N., Avis, D., Evans, A. C., 2000. Automated 3-d extraction of inner and outer surfaces of cerebral cortex from mri. NeuroImage 12 (3), 340–356.

Mathotaarachchi, S., Pascoal, T. A., Shin, M., Benedet, A. L., Kang, M. S., Beaudry, T., Fonov, V. S., Gauthier, S., Rosa-Neto, P., Initiative, A. D. N., et al., 2017. Identifying incipient dementia individuals using machine learning and amyloid imaging. Neurobiology of Aging.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., et al., 2011. The diagnosis of dementia due to alzheimers disease: Recommendations from the national institute on aging-alzheimers association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimer's & dementia 7 (3), 263–269.

Misra, C., Fan, Y., Davatzikos, C., 2009. Baseline and longitudinal patterns of brain atrophy in mci patients, and their use in prediction of short-term conversion to ad: results from adni. Neuroimage 44 (4), 1415–1422.

Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., Hodges, J. R., Sep. 2009. Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm. Journal of neurology 256 (9), 1500–1509.
URL http://dx.doi.org/10.1007/s00415-009-5152-0

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., Initiative, A. D. N., et al., 2015. Machine learning framework for early mri-based alzheimer's conversion prediction in mci subjects. Neuroimage 104, 398–412.

Orban, P., Tam, A., Urchs, S., Savard, M., Madjar, C., Badhwar, A., Dansereau, C., Vogel, J., Shmuel, A., Dagher, A., Villeneuve, S., Poirier, J., Rosa-Neto, P., Breitner, J., Bellec, P., , , 2017. Subtypes of functional brain connectivity as early markers of neurodegeneration in alzheimer's disease. bioRxiv.
URL https://www.biorxiv.org/content/early/2017/09/28/195164

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., Fratiglioni, L., 2014. Mild cognitive impairment: a concept in evolution. Journal of internal medicine 275 (3), 214–228.

Phipson, B., Smyth, G. K., 2010. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. Statistical applications in genetics and molecular biology 9 (1).

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Feb. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59 (3), 2142–2154.
URL http://dx.doi.org/10.1016/j.neuroimage.2011.10.018

Querbes, O., Aubry, F., Pariente, J., Lotterie, J.-A., Démonet, J.-F., Duret, V., Puel, M., Berry, I., Fort, J.-C., Celsis, P., et al., 2009. Early diagnosis of alzheimer's disease using cortical thickness: impact of cognitive reserve. Brain 132 (8), 2036–2047.

Rabinovici, G. D., Carrillo, M. C., Forman, M., DeSanti, S., Miller, D. S., Kozauer, N., Petersen, R. C., Randolph, C., Knopman, D. S., Smith, E. E., et al., 2017. Multiple comorbid neuropathologies in the setting of alzheimer's disease neuropathology and implications for drug development. Alzheimer's & Dementia: Translational Research & Clinical Interventions 3 (1), 83–91.

Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., Davatzikos, C., 2017. A review on neuroimaging-based classification studies and associated feature extraction methods for alzheimer's disease and its prodromal stages. NeuroImage.

Sled, J. G., Zijdenbos, A. P., Evans, A. C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in mri data. IEEE transactions on medical imaging 17 (1), 87–97.

Smith, S. M., Nov. 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155.
URL http://dx.doi.org/10.1002/hbm.10062

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., Wagster, M. V., Phelps, C. H., May 2011. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association

workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7 (3), 280–292.
URL http://dx.doi.org/10.1016/j.jalz.2011.03.003

Stephan, B., Hunter, S., Harris, D., Llewellyn, D., Siervo, M., Matthews, F., Brayne, C., 2012. The neuropathological profile of mild cognitive impairment (mci): a systematic review. Molecular psychiatry 17 (11), 1056–1076.

Tohka, J., Zijdenbos, A., Evans, A., 2004. Fast and robust parameter estimation for statistical partial volume models in brain mri. Neuroimage 23 (1), 84–97.

Urchs, S., Armoza, J., Benhajali, Y., St-Aubin, J., Orban, P., Bellec, P., 2017. Mist: A multi-resolution parcellation of functional brain networks [version 1; referees: awaiting peer review]. MNI Open Research 1 (3).

Walt, S. v. d., Colbert, S. C., Varoquaux, G., 2011. The numpy array: a structure for efficient numerical computation. Computing in Science & Engineering 13 (2), 22–30.

Wee, C.-Y., Yap, P.-T., Shen, D., 2013. Prediction of alzheimer's disease and mild cognitive impairment using cortical morphological patterns. Human brain mapping 34 (12), 3411–3425.

Woo, C.-W., Chang, L. J., Lindquist, M. A., Wager, T. D., 2017. Building better biomarkers: brain models in translational neuroimaging. Nature neuroscience 20 (3), 365–377.

Xu, L., Wu, X., Chen, K., Yao, L., 2015. Multi-modality sparse representation-based classification for alzheimer's disease and mild cognitive impairment. Computer methods and programs in biomedicine 122 (2), 182–190.

Zhu, X., Suk, H.-I., Shen, D., 2014. A novel matrix-similarity based loss function for joint regression and classification in ad diagnosis. NeuroImage 100, 91–105.

Zu, C., Jie, B., Liu, M., Chen, S., Shen, D., Zhang, D., Initiative, A. D. N., et al., 2016. Label-aligned multi-task feature learning for multimodal classification of alzheimers disease and mild cognitive impairment. Brain imaging and behavior 10 (4), 1148–1159.

Supplementary Material – A brain signature highly predictive of future progression to Alzheimer's dementia

Table S2: Performance of the models. Prec: precision, Spec: specificity, Sens: sensitivity and N: number of selected subjects.

Modality	Algo	Contrast	Prec (%)	Spec (%)	Sens (%)	N
fMRI	Base	CN/AD	38.10	46.94	66.67	42
	HPS		60	95.92	12.5	5
sMRI	Base		66.67	83.67	66.67	24
	HPS		87.50	97.96	29.17	8
fMRI+sMRI	Base		69.57	85.71	66.67	23
	HPS		100	100	37.50	9
fMRI+sMRI	Base	sMCI/pMCI	73.33	89.19	57.89	15
	HPS		90	97.3	47.37	10

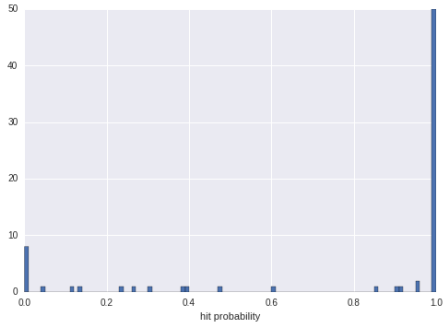

Figure S1: Hit-probability distribution obtained from replicating the SVM training 100 times from 80% of the training set.